



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 163761

TO: Patricia Duffy
Location: 3b05 / 3c18
Sunday, August 28, 2005
Art Unit: 1645
Phone: 571-272-0855
Serial Number: 09 / 868195

From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504
jan.delaval@uspto.gov

Search Notes

This Page Blank (uspto)

From: Duffy, Patricia
Sent: Thursday, August 25, 2005 8:33 AM
To: STIC-Biotech/ChemLib
Subject: sequence search 09/868195

In re: 09/868,195

Please search SEQ ID NO:12.

Please search commercial and interference databases.

Please print out top 75 hits in each category.

Thank you.

Patricia A. Duffy, Ph.D.
Art Unit 1645
Remsen 3B05; Mailbox 3C18
571-272-0855

STAFF USE ONLY

Searcher: Jan
Searcher Phone: 2- 2504
Date Searcher Picked up: 8/28/05
Date Completed: 8/28/05
Searcher Prep/Rev. Time: 0
Online Time: 12

Type of Search

NA#: _____ AA#: ✓
Interference: ✓ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: ✓
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2005, 10:18:29 ; Search time 41 Seconds

(without alignments)
1114.706 Million cell updates/sec

Title: US-09-868-195-12

Perfect score: 2393
Sequence: 1 LRKEYQNYNGEMKSVNOI.....GIRYSIEANTNKSYLDK 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 80 summaries

Database :

1: PIR 79: **
2: pir2: **
3: pir3: **
4: pir4: **

Pred. No. is the number of results predicted by change to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1812	75.7	474	2	E95129
2	1802	75.3	474	2	D98000
3	1702	71.1	475	2	A57151
4	1383	57.8	482	2	H97348
5	1244	52.0	481	2	E83929
6	1168.5	48.8	496	2	F84634
7	1164.5	48.7	498	1	S43833
8	1143.5	47.8	496	1	S43832
9	868.5	36.3	475	2	F82900
10	763	31.9	488	2	G69752
11	753.5	31.5	463	2	B64476
12	752	31.4	470	2	F90347
13	750	31.3	470	2	D90323
14	741.5	31.0	455	2	H69230
15	741.5	31.0	489	2	A83033
16	737	30.8	476	2	A70318
17	736.5	30.8	491	2	T46684
18	735	30.7	482	2	B83667
19	726	30.3	482	2	F65045
20	721.5	30.2	488	1	I39769
21	720	30.1	488	2	A11546
22	718	30.0	488	2	A11188
23	717	30.0	482	2	D85913
24	713.5	29.8	482	2	B91669
25	713.5	29.8	480	2	C84250
26	710.5	29.7	489	2	T44939
27	710.5	29.7	489	2	C84401
28	710.5	29.7	498	2	AD2377
29	697.5	29.1	468	2	C83774

30	682.5	28.5	485	2	B83901	aldehyde dehydroge
31	682	28.5	482	2	A80839	succinate-semialde
32	682	28.5	496	2	T44987	aldehyde dehydroge
33	673	28.1	485	2	C69584	aldehyde dehydroge
34	669.5	28.0	470	2	A87547	vanillin dehydroge
35	668	27.9	483	2	D83613	succinate-semialde
36	666.5	27.9	524	2	E75574	aldehyde dehydroge
37	666	27.8	509	1	S14629	aldehyde dehydroge
38	664.5	27.8	495	2	C89778	hypothetical prote
39	664	27.7	477	2	A83469	aldehyde dehydroge
40	661.5	27.6	509	2	E90504	hypothetical prote
41	661.5	27.6	509	2	E96825	hypothetical prote
42	661	27.6	517	1	A40872	aldehyde dehydroge
43	660	27.6	483	2	E49343	salicylaldehyde de
44	658.5	27.5	490	1	S1181	betaine-aldehyde d
45	658	27.5	501	1	DEH0E1	aldehyde dehydroge
46	656	27.4	485	2	C95964	probable aldehyde
47	655	27.4	487	2	AD3300	succinate-semialde
48	653.5	27.3	490	2	B85524	NAD--dependent bet
49	653	27.3	477	2	A95990	probable dehydroge
50	651.5	27.2	485	2	H72562	probable aldehyde
51	651.5	27.2	487	2	G75592	succinate-semialde
52	650	27.2	490	2	E95316	aldehyde dehydroge
53	650	27.2	512	1	A55684	aldehyde dehydroge
54	649.5	27.1	490	2	AD0143	betaine aldehyde d
55	649	27.1	496	2	T50272	probable aldehyde
56	649	27.1	996	2	AC1874	1-pyrroline-5 carb
57	648.5	27.1	490	2	F90673	NAD--dependent bet
58	647.5	27.1	495	2	A75608	aldehyde dehydroge
59	647	27.0	495	1	G64878	aldehyde dehydroge
60	644.5	26.9	462	2	D69764	succinate-semialde
61	643.5	26.9	485	2	C69814	benzaldehyde dehyd
62	643.5	26.9	490	2	C83675	glycine betaine al
63	643	26.9	495	2	E90863	aldehyde dehydroge
64	643	26.9	495	2	D85755	aldehyde dehydroge
65	640	26.7	462	2	T36807	probable aldehyde
66	639	26.7	484	2	A83176	NAD-dependent succ
67	639	26.7	501	2	S14752	aldehyde dehydroge
68	639	26.7	990	2	S75910	probable proline d
69	637.5	26.6	485	2	E83715	aldehyde dehydroge
70	637.5	26.6	486	2	T36346	probable aldehyde
71	637	26.6	479	2	T13293	aldehyde dehydroge
72	636	26.6	484	2	F98208	succinate-semialde
73	636	26.6	484	2	AC3078	aldehyde dehydroge
74	635	26.5	499	2	S74224	aldehyde dehydroge
75	632.5	26.4	483	2	H98159	hypothetical prote
76	632.5	26.4	483	2	A13127	vanillin: NAD oxid
77	630.5	26.3	497	1	A29055	aldehyde dehydroge
78	630	26.3	497	2	B82981	probable aldehyde
79	628.5	26.3	480	2	A81023	aldehyde dehydroge
80	627.5	26.2	480	2	F81965	lactaldehyde dehyd

ALIGNMENTS

RESULT 1

hypothetical protein sp119 [imported] - Streptococcus pneumoniae (strain TIGR4)

C/Spectrum: Streptococcus pneumoniae

C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Jul-2004

C/Accession: E95129

R/Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Umayam, L.A.; White, O.; Salberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,

neon, T.; Hickey, B.K.; Holt, I.E.

Science 293, 498-506, 2001

A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A/Reference number: A95000; MUID:21357209; PMID:11463916

A/Accession: E95129

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-474 <KUR>

A:Cross-references: UNIPROT:Q97Q70; GB:AE005672; PIDN:AAK75230.1; PID:g14972596; GSPDB:G
A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP119

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 75.7%; Score 1812; DB 2; Length 474;

Best Local Similarity 74.1%; Pred. No. 7.6e-108;

Matches 349; Conservative 57; Mismatches 65; Indels 0; Gaps 0;

```
QY 5 YQNVNGEKSSVNOQIEILSPIDSSLGVPAMTRREVDHAKAGREALPAMALTVER 64
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 YQNLVNGKWKSSQEOETIYSPINOEBELGTVPAMTQTEADEAMQAAAPALPARALSAVER 63
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 65 AQLYHKAADIIRDKKEIATVLAKESIKAVNASVTEVVRTADLIRYAAEBGIRLSTSADE 124
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 AAYLHKAALIERDKKEIGTILLAKEVAKGIKAAIGEVVRTADLIRYAAEBGIRLTQQAHE 123
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 125 GGRMDASTGKHLAVIRQPVGIVLAIAPYNPVNLSGSKIAPALIGNVNMFKEPPTQSGV 184
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 GGGFEATSKNKLAVRREPVGIVLAIAPFNYPVNLASAKIAPALIGNVNMFKEPPTQSGI 183
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 185 SGLVLAKAPAEAGLPAGVFNITTRGSGEIGDYIIEHEVNFNFTGSPVQORIGKLAGM 244
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 SGLLAKAFEEBAGIPAGVFNITTRGSGEIGDYIIEHEVNFNFTGSPVQORIGKLAGM 243
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 245 RPIMLEGKDGAGIVLADADLDNAAKOIVAGAYDVSQGRCTAIRVLYEEVADLAEKI 304
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 RPIMLEGKDALVLEADLEHAAKOIVAGAFSTSGQRCTAIRVLYESVADKLATLL 303
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 305 SENVAKLSVGDPEFDNATVPVIIDNSADPIESLVVDARQKAKELNEFKRDGRLLTPGLF 364
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 304 QEEVSKLTVGDPFDNADITPVIDNASADPIWGLIEDAQEKAQALTPIKREGNLLMPVLF 363
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 365 DHVTLDMKLAMEBPPGPIIPIIRYKDAEAVAIANKSDGLQSSVTRPFOKAFDIANKL 424
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 364 DQVTKDMKVAMEBPPGPVLPPIIRVASVEBALAFANESBFGLOSSVPTNDFKKAFAIEKEL 423
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 425 EYGVTHINNKTGRGPDNPFPLGLKSGAGVQIGIRYSIEMTNVKSIVLDMK 475
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 424 EYGVTHINNKTGRGPDNPFPLGVKSGAGVQIGIKYSIEMTNVKSIVPDVK 474
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 2

D98000

Glyceraledehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.9) gapN [imported] - Streptococcus

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C:Accession: D98000

R:Hostins: J.A.; Albhorn Jr., W.; Arnold, J.; Biaszcak, L.; Burgett, S.; Dehoff, B.S.; F

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellio, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: D98000

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-474 <KUR>

A:Cross-references: UNIPROT:Q8DP87; GB:AE007317; PIDN:AAK99832.1; PID:g15458646; GSPDB:G

C:Genetics:

A:Gene: gapN

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

C:Keywords: oxidoreductase

Query Match 75.3%; Score 1802; DB 2; Length 474;

Best Local Similarity 73.7%; Pred. No. 3.3e-107;

Matches 347; Conservative 58; Mismatches 66; Indels 0; Gaps 0;

```
QY 5 YQNVNGEKSSVNOQIEILSPIDSSLGVPAMTRREVDHAKAGREALPAMALTVER 64
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 YQNLVNGKWKSSQEOETIYSPINOEBELGTVPAMTQTEADEAMQAAAPALPARALSAIER 63
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
QY 65 AQLYHKAADIIRDKKEIATVLAKESIKAVNASVTEVVRTADLIRYAAEBGIRLSTSADE 124
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 AAYLHKAALIERDKKEIGTILLAKEVAKGIKAAIGEVVRTADLIRYAAEBGIRLTQQAHE 123
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 125 GGRMDASTGKHLAVIRQPVGIVLAIAPYNPVNLSGSKIAPALIGNVNMFKEPPTQSGV 184
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 GGGFEATSKNKLAVRREPVGIVLAIAPFNYPVNLASAKIAPALIGNVNMFKEPPTQSGI 183
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 185 SGLVLAKAPAEAGLPAGVFNITTRGSGEIGDYIIEHEVNFNFTGSPVQORIGKLAGM 244
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 SGLLAKAFEEBAGIPAGVFNITTRGSGEIGDYIIEHEVNFNFTGSPVQORIGKLAGM 243
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 245 RPIMLEGKDGAGIVLADADLDNAAKOIVAGAYDVSQGRCTAIRVLYEEVADLAEKI 304
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 RPIMLEGKDALVLEADLEHAAKOIVAGAFSTSGQRCTAIRVLYESVADKLATLL 303
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 305 SENVAKLSVGDPEFDNATVPVIIDNSADPIESLVVDARQKAKELNEFKRDGRLLTPGLF 364
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 304 QEEVSKLTVGDPFDNADITPVIDNASADPIWGLIEDAQEKAQALTPIKREGNLLMPVLF 363
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 365 DHVTLDMKLAMEBPPGPIIPIIRYKDAEAVAIANKSDGLQSSVTRPFOKAFDIANKL 424
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 364 DQVTKDMKVAMEBPPGPVLPPIIRVASVEBALAFANESBFGLOSSVPTNDFKKAFAIEKEL 423
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 425 EYGVTHINNKTGRGPDNPFPLGLKSGAGVQIGIRYSIEMTNVKSIVLDMK 475
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 424 EYGVTHINNKTGRGPDNPFPLGVKSGAGVQIGIKYSIEMTNVKSIVPDVK 474
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 3

A57151

Glyceraledehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.9) - Streptococcus mutans

C:Species: Streptococcus mutans

C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 09-Jul-2004

C:Accession: A57151

R:Boyd, D.A.; Cvitkovitch, D.G.; Hamilton, I.R.

J. Bacteriol. 177, 2622-2727, 1995

A:Title: Sequence, expression, and function of the gene for the nonphosphorylating, NADP

A:Reference number: A57151; MUID:95270576; PMID:7751269

A:Accession: A57151

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-475 <BOY>

A:Cross-references: UNIPROT:Q59931; GB:IJ38521; GB:IJ5191; NID:g642666; PIDN:AAA91091.1;

C:Genetics:

A:Gene: gapN

A:Start codon: TTG

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

C:Keywords: NADP; oxidoreductase

F:41-304/Domain: aldehyde dehydrogenase homology <AND>

Query Match 71.1%; Score 1702; DB 2; Length 475;

Best Local Similarity 66.9%; Pred. No. 7.6e-101;

Matches 318; Conservative 75; Mismatches 82; Indels 0; Gaps 0;

```
QY 1 LTKYQNVNGEKSSVNOQIEILSPIDSSLGVPAMTRREVDHAKAGREALPAMALT 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MTKQYKNVNGEKSSQENIKIYEPASGAELGSPVAMTEEDVYASAKAQAPMARLS 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 YVERAQYHKAADIIRDKKEIATVLAKESIKAVNASVTEVVRTADLIRYAAEBGIRLST 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 YIERAAIYHKAADILMRKKEIKGALISKVAKGYNSVSVTRTAEIINTYAAEBGLRMEG 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 SADBGKMDASTGKHLAVIRQPVGIVLAIAPYNPVNLSGSKIAPALIGNVNMFKEPPT 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 EYLGGSPFAASKKIALAVRREPVGIVLAIAPFNYPVNLASAKIAPALIGNVNMFKEPPT 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 QGSGVGLTAAFAAGLPAGVFNITTRGSGEIGDYIIEHEVNFNFTGSPVQORIGK 240
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 QGISGGLLAEFAAGLPAGVFNITTRGSGEIGDYIIEHQAQVNFNFTGSGEIGERIGK 240
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 LAGMRPIMLEGKDGAGIVLADADLDNAAKOIVAGAYDVSQGRCTAIRVLYEEVADL 300
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


Db 312 DALVEKVKVAKLSTVPEEDSDITPVVSESSANFLBGLVMDAKEXKATPCQEVYKREGN 371
QY 358 LITPGFEDHYTLDMKLAWEPPFPIPIIRVKAEEAVAIANKSDPGLQSSVTRPQKA 417
Db 372 LIMPILLDNRPRMIRAMEBPFQVPIVIRINSVEBGIHHNCNANSGLOQCVCTTKDINRA 431
QY 418 FDIANKLEVTGVIINNKTKGRGPNPPLGLKSGAGVQIGIRYSIEAMTWKSVI LDM 474
Db 432 IMISDMESGTVQINSAPARGPHFSPQIGKDSIGSQGITSINMMTKVITVNL 488

RESULT 9
F82900
NADP-dependent glyceraldehyde-3-phosphate dehydrogenase U0362 [imported] - Ureaplasma urealyticum
C/Species: Ureaplasma urealyticum
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Mar-2004
C/Accession: F82900
R/Glaes, J.I.; Lefkowitz, E.J.; Glaes, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mir
A/Reference number: A82870
A/Accession: F82900
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-475 <GLA>
A/Cross-references: GB:AE002133; GB:AF222894; NID:g66939339; PIDN:AAF30771.1; GSPDB:GN001
A/Experimental source: serovar 3; biovar 1
A/Genetic: gapN; U0362
A/Genetic code: SGC3
C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 36.3%; Score 868.5; DB 2; Length 475;
Best Local Similarity 39.7%; Pred. No. 8.4e-48;
Matches 190; Conservative 95; Mismatches 177; Indels 17; Gaps 5;

QY 5 YONVYNGEKKSSVNOJELISPIDSSLGFPVPMATRE-EVDHMKAGREALPAMALTYVE 63
Db 3 YKLINGAFVDAEKKLPVNPSPNOQIAVAPNHNENEINTTFENAHIVLYKPKDPIRX 62
QY 64 RAQYIKAAIDIERDEKEIATVLAKEISKAYNASVTEVTRADLIRYAAEGIRLSTSA- 122
Db 63 RCDLLKLADKLDENHQLAQIISTEIAQLKDSLEVERSDYLAETVFEYQKLMQKRI 122
QY 123 --DEGKMDASTGH---KLAIVTRQPGVIGVIAIADYNTPVNLSGSKIAPALIGNVMP 176
Db 123 IFDE-----TVHAYKNKVAATYRIPVGVVLAICFPNYPINLLISKLPALVSGNSLVY 175
QY 177 KPTQGSVSGVILAKAFAGLPAQVFNITIGSGSEIGYIVHEEVNFINTGSTPVGQ 236
Db 176 KSTQGSLLGIRISELVHEGPKGVNCLTTEARITGDLVTNKVKVKAISFTGGPRVN 235
QY 237 RIGKLAGMPIMLELGKDAAGVILADADLDNNAKOIVAGAYDSGQRTAIRVLAIVEY 296
Db 236 HIAETISKISLVHELGGOPALVLDADBELANETVKAIGSGQRCTAIRVAFVSHQK 295
QY 297 ADELAEKISENVAKLSVGDPPDNATVPIVDNSADPISLVVDARQKA--KEINLFRK 354
Db 296 HDLVLVILNKKVDALTVGLPQGNPITPLINSNLKYNLSLVEDAIKKGAIVHQKLVNE 355
QY 355 DGRLLITPGFEDHYTLDMKLAWEPPFPIPIIRVKAEEAVAIANKSDPGLQSSVTRP 414
Db 356 KNNLHPVLDVNTKMRVAMEBPFPIPIITVNSIQEALIDINSQYGLQACIPTTVY 415
QY 415 OKAFDIANKLEVTGVIINNKTKGRGPNPPLGLKSGAGVQIGIRYSIEAMTWKSVI LDM 473
Db 416 ASTEQALQIESGTININKSSSRGPDILPFGVKGDFGVQGLVDAILSMITIKGIIN 474

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: G69752
R/Kunze, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C./Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A./Enllich, S.D.; Emerson, P.T.; Enlian, K.D.; Erington, J.; Fabret, C.; Ferrati, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallie
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F
Koeltter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Maueel
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekoska, A.; Seru
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Teppstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumeito, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: G69752
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-488 <KUN>
A/Cross-references: UNIPROT:P42236; GB:Z99105; GB:AL009126; NID:g2632457; PIDN:CAB12041
A/Experimental source: strain 168
C/Genetics:
A/Genetic: ycbd
C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
F/49-309/Domain: aldehyde dehydrogenase homology <ALD>

Query Match 31.9%; Score 763; DB 2; Length 488;
Best Local Similarity 36.9%; Pred. No. 4.5e-41;
Matches 170; Conservative 94; Mismatches 179; Indels 18; Gaps 6;

QY 5 YONVYNGE--KSSVNOJELISPID-SSLGFPVPMATREEDVHMKAGREALPAMALTYV 61
Db 10 YNLPINGEYVKSQSDMVKVNPVNDIVGVQNSTADVERAYTAANEATARKLTLG 69
QY 62 YRRAQYIKAAIDIERDEKEIATVLAKEISKAYNASVTEVTRADLIRYAAEGIRLST 121
Db 70 AERGOVLYKTADIMEORLEIIAACATREWGKTLPEAKGETARGLAIRYAAEGGRKT-- 127
QY 122 ADEGKMDASTGHKLAIVTRQPGVIGVIAIADYNTPVNLSGSKIAPALIGNVMPKPPQ 181
Db 128 ---GDVISTDQDALMFTTRVPLVAVGVISPNPFALPYMKAPALVAGNVVLIKPAE 184
QY 182 GSVSGVILAKAFAGLPAQVFNITIGSGSEIGYIVHEEVNFINTGSTPVGQRIKT 241
Db 185 TAVTCAKIIACFEBAGLPAGVINLVTPGCSVVGGLAHDGVNAVTFGTSNQVKIIGA 244
QY 242 AQMR--PIMLELGKDAAGVILADADLDNNAKOIVAGAYDSGQRTAIRVLAIVEY 299
Db 245 ALARGAKYOLEMGKGNPVIVADADLEAABAVIIGAFSTQCKTATSRVIVQSGIYER 304
QY 300 LAEKISENVAKLSVGDPP--DNATVPIVDNSADPISLVVDARQKAEL-----NE 351
Db 305 FKEKILQRTKDTITDLSKEDVMMGPILASKNDLCLSTIEKQKOGASLLIGEKLENG 364
QY 352 FRDGRLLITPGFEDHYTLDMKLAWEPPFPIPIIRVKAEEAVAIANKSDPGLQSSVPT 411
Db 365 KYONGVYVQPAFDVNTSEMTIAQSEIFQPVIALIKVDSIEALMIANDVKRGLASIT 424
QY 412 RDKAFDIANKLEVTGVIINNKTKGRGPNPPLGLKSGAG 452
Db 425 ENIGMLSFIDEIDAGLVINAESAGVELQAFPGMKOSSS 465

RESULT 11
B64476
NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.-) - *Methanococcus jan*
C/Species: *Methanococcus jannaschii*
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C/Accession: B64476
R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klank, H.P.; Fraser, C.M.; Smith, H.O.; Moese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MIMD:96337939; PMID:8688087

A:Accession: B64476

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-463 <BUL>

A:Cross-references: UNIPROT:Q58806; GB:U67581; GB:L77117; NID:g1592052; PID:g1592060; TI

C:Genetics:

A:Map position: FOR371808-1373199

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

C:Keywords: NAD; oxidoreductase

Query Match 31.5%; Score 753.5; DB 2; Length 463;

Best Local Similarity 36.9%; Pred. No. 1.7e-40;

Matches 174; Conservative 95; Mismatches 184; Indels 19; Gaps 8;

8 YNNGEWSVNO--IEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVERA 65

2 FIDGK---INEMDMVINPYSLEVKTIPLSREAKKALIDAEKYEVMKRLPTTKY 58

66 QYIHKRAADIEERDKKEIATVLAKESKAYNASVTEVTRADLIRYAAEEGRLSTADG 125

59 NIIIMINAKQIKERKEELAKILADAGKPIKQARVEVERSIGTFKLAA--FYVEHRDEV 115

126 GKGDASTGHKLAIVIRQPGIVLALAPYVPVNLSSSKIAIPALIGNVMPFPPOGAS 185

116 IRSD---DRLIFTRKEPVGIVALTIPNPLNSHKIAPALATGNVIVHPSKAPLV 171

186 GLVLA---APAEAGLPAGVENTITGRGSEIGDYVEHEEVNFIPTGSTPGORIGL 241

172 CIELAKIINLAKKYNVPGLVNLTGAGEVVDDELVENKNNMISFTSSKVGELITTK 231

242 AGRRPIMLELGGKAGIVLADADLDAKQIVAGAYDSGORCTAIKRVLYVEVADDEL 301

232 AGFKKIALTELGGVNPVIVLKDADLNKAVNALIKGSPYVAGQVCSIGMILVDESINDK 291

302 EKISENVALLSGDPPDNAT-VTPVIDNSADPIESLVVDAROKAKELNEFRDRLIT 360

292 EHFVNKAKVNVGNPDEKTDVGPLISVBAHEWVEVKKALIDEGSKLLGGRRDKALFY 351

361 PGLFDHTLDMKLAWEPPPGIPIIRVDAEBAVAIANKSDPGLQSSVTRPFOKAPDI 420

352 PTLLE-VDDNLTICKETETAPVPIIRTRME-EEIMIANSTETGLSALFTINDINKSLAF 409

421 ANKLEVGTVHNNKTGRGPDNPFPLGLKSGGAGVGIRYSIEAMTNVKSIVL 472

410 AENLEFGVAVINDSLFRQDNMFPFGVKKSGLGREGVKTAMEEMSNIXTII 461

Db

RESULT 12

P90347

hypothetical protein gapN-2 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: P90347

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F

arrett, R.A.; Regan, M.A.; Sersen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: P90347

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-470 <KUR>

A:Cross-references: UNIPROT:Q97XA5; GB:AE006641; NID:g13815099; PIDN:AAK42037.1; GSPDB:C

C:Genetics:

A:Gene: gapN-2

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 31.4%; Score 752; DB 2; Length 470;

Best Local Similarity 36.5%; Pred. No. 2.2e-40;

Matches 172; Conservative 96; Mismatches 187; Indels 16; Gaps 7;

16 SVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVERAQLHKAADI 75

3 LIEIVRSPNSLTKIVGTRKMSKDEVGIEBAVKGEIISRMPLVKTALIRKISEIL 62

76 ERDKKEIATVLAKESKAYNASVTEVTRADLIRYAAEE-GIRLSTADBEGRKMDA---- 130

63 EREGRRLARLAMEGKPISDSRVEVLASRLFRPAAEVGVLE--GKNRVDAVAYTP 119

131 -STGKLAIVIRQPGIVLALAPYVPVNLSSSKIAIPALIGNVMPFPPOGASVGLV 189

120 PENEMRIYLSREPVGIVTALIPFPFINSRAHKVAPALVAGNSVVPFINTPLAAVEM 179

190 AKAPAEAGLPAGVENTITGRGSEIGDYVEHEEVNFIPTGSTPGORIGKLA--GNRP 246

180 KKIIVEAGLPDSAVVTVVYGSSEIGDEIITHPVLGLITLTGSTGTGLASKALSIGR- 238

247 IMLLEIGKDAIVLADADLDAKQIVAGAYDSGORCTAIKRVLYVEVADDELAKISE 306

239 IIMELGSSDPITILEDANIERASSIAVARPEVAGQNCNAGKRITIVROEVTDKVKAFNE 298

307 NVAKLSDVPDPDNAT-VTPVIDNSADPIESLVVDAROKAKE--LNEFKDGRLLTPGL 363

299 KAKALKVGEPELDTTDVGPVINKSEVENLNVLEDAKVKGRVLEILNKGSSGSPFLTM 358

364 FDHTLDMKLAWEPPPGIPIIRVDAEBAVAIANKSDPGLQSSVTRPFOKAPDIANK 423

359 VTPSLDMLVYKSEVFPGIVPVSYSKDEFAIRINSTREYGOSAFINNDVRAKLKSRE 418

424 LEVGTVHNNKTGRGPDNPFPLGLKSGGAGVGIRYSIEAMTNVKSIVLDM 474

419 LKFGAVIINDSTRLRMSDLPFGGFKTGIIRGBVARETMELENTENKLIATTL 469

Db

RESULT 13

D90323

hypothetical protein gapN-1 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: D90323

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F

arrett, R.A.; Regan, M.A.; Sersen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: D90323

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-470 <KUR>

A:Cross-references: UNIPROT:Q97XS9; GB:AE006641; NID:g13814871; PIDN:AAK41843.1; GSPDB:C

C:Genetics:

A:Gene: gapN-1

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 31.3%; Score 750; DB 2; Length 470;

Best Local Similarity 36.6%; Pred. No. 2.9e-40;

Matches 171; Conservative 94; Mismatches 186; Indels 16; Gaps 7;

20 IEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVERAQLHKAADIEERDK 79

7 IEVRSPNMLNVIGTVKRNHKOEVREIEBAVKGEIISKMPYKKTALIRKSEILERBQ 66

80 EEIATVLAKESKAYNASVTEVTRADLIRYAAEE-GIRLSTADBEGRKMDA----STG 133

67 EELATLAMEAGKPIRDSRVEVLASRLFRHAAEEGVNLE--GKATRVDAVAYTPPNE 123

134 HKLAIVIRQPGIVLALAPYVPVNLSSSKIAIPALIGNVMPFPPOGASVGLVLAAR 193

Db 124 NRIVSTREPIGVVTAALPFNPFLINSPAHKAVADALVAGNSVVKPSINTPLAITEMKIL 183
Qy 194 AEEGLPAGVNTTGGSEIGDYIVHEEVPNFPGSTVQRIQGLA--GMRPLME 250
Db 184 VENGPLPSAVRVYTGSSSEIGDEIITHTPLVGLITLGTGTGKIKASKAVSLKR-1IME 242
Qy 251 LGGKDAQIVLADDLNAAKOIVAGAYDSGQCTAIKRVLYVEEVDDELAEKISENVAK 310
Db 243 LGGSDPFIILEDNANIEASSIAVRARFEIYAGQNCNCKRIIEBEIDYKAKAFADKARA 302
Qy 311 LSVGDPEDNAT-VTPVIDNSADFIESTLVVDAROKAKE--LNEFKRGRLLTPGFLDHY 367
Db 303 LKVGDPLEDETVDGPVINKSEVENINSVLADAKAGKVEILNKGPSSGFFPLTWNTNP 362
Qy 368 TLDMKLAMEBPPFGPIPLIIRVYKDAEVAIAJANKSDFGQSVTRDFOKAFDIANKIEVG 427
Db 363 SLMLVLKSEVFPPIVIVSVKDEEAIKIANSTEVGLQGAIFPDNDVRALKLSRELKFG 422
Qy 428 TVHINKTGGRPNPFPLGKSGAGVQGIKRISEIEMTNKSIYLDK 474
Db 423 AVIINSTRLRMDSLAFGFGFKSSIGREVTLEMTENKLIATIL 469

RESULT 14

H69230
NADP-dependent glyceraldehyde-3-phosphate dehydrogenase - Methanobacterium thermoautotrophicum
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: H69230
R/Smith, D.R.; Doucette-Stamm, L.A.; Delouche, C.; Lee, H.; Dubois, J.; Alredge, T.;
Qiu, D.; Spadecora, R.; Vitacore, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
Kt. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A/Reference number: A69000; MUID:98037514; PMID:9371463
A/Accession: H69230
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-455 <MTH>
A/Cross-references: UNIPROT:Q27059; GB:AE000871; GB:AE000666; MID:92622069; PIDN:AA8547
A/Experimental source: strain Delta H
C/Genetics:
A/Genes: MTH978
C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 31.0%; Score 741.5; DB 2; Length 455;
Best Local Similarity 38.2%; Pred. No. 9,6e-40;
Matches 173; Conservative 83; Mismatches 186; Indels 11; Gaps 5;
Qy 9 VNGEMKSVNQIILSPIDSSLGFPVPMTRREVDHAKAGREALPMAALTYVERAQL 68
Db 5 IDEGVSGEKIFVTRNPFNGDEVDVPLAGRSVDERAIRAHRARADAMADLSARKISEKL 64
Qy 69 HKRAADIIRKKEIATVLAKEISKAYNAVSTEVRADLIRYAEGRILSTADSGKM 128
Db 65 YVVADELKTELDFARILTESGKPIRESDEKRSVETARLSAEGR--YGSIPM 121
Qy 129 DASTGKH--LAVRROPVGIYLAIPYNPVNLSGSKIAPALIGNVMMKPPQGSVSG 186
Db 122 DAIIGKGLGTFVRIPLGVVAITPPNYPLNLAHKVGRPALAAGTTTIKPSLEAPLSA 181
Qy 187 LVYAKPAEAGLPAGVNTTGGSEIGDYIVHEEVPNFPGSTVQRIQGLAGMRP 246
Db 182 LKIAMILISE-HFPAAGVNAVATGSGSEVGVIIIDSPVDKITTGSEVGVYISARASMK 240
Qy 247 IMLELGGKAGIYLAADDLNAAKOIVAGAYDSGQCTAIKRVLYVEEVDDELAEKISE 306
Db 241 ITBELGNDPLIYVDDADISAVEAIVKSYLSGQCLIAVHKAIYHEDVADFPADLVN 300
Qy 307 NVAKLSVGDPPDNAT-VTPVIDNSADFIESTLVVDAROKAKEINERKGRLLTPGLPD 365
Db 301 ITGSLRAGBMDVTRVDVGLINDEAIEVERVIGAAVEDGAEILLCGSGRGNFVEPTVLD 360

Qy 366 HTVLDMKLAMEBPPFGPIPLIIRVYKDAEVAIAJANKSDFGQSVTRDFOKAFDIANKIE 425
Db 361 HVPKMEVIERETFGVSVIIRNENGADEAIRANGCTYALQAGVFENIRITALMARREIE 420
Qy 426 VGTVHINKTGGRPNPFPLGKSGAGVQGIKRISEIEMTNKSIYLDK 474
Db 421 AGTVLVNKGSTFRVDMFPG----GDGAVBWR 449

RESULT 15

A83033
Probable aldehyde dehydrogenase PA4899 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: A83033
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: A83033
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-489 <STO>
A/Cross-references: UNIPROT:Q9HUR4; GB:AE004903; GB:AE004091; MID:93951173; PIDN:AA60828
A/Experimental source: strain PA01
C/Genetics:
A/Genes: PA4899
C/Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 31.0%; Score 741.5; DB 2; Length 489;
Best Local Similarity 37.0%; Pred. No. 1.1e-39;
Matches 173; Conservative 87; Mismatches 185; Indels 23; Gaps 7;

Qy 11 GEMK--SSVNOIILSPIDSSLGFPVPMTRREVDHAKAGREALPMAALTYVERAQL 68
Db 16 GEMHGRAGRRLVNSPFGSLLEIETQADRDLDLAIVKAAEVQPMALGPSARAYL 75
Qy 69 HKRAADIIRKKEIATVLAKEISKAYNAVSTEVRADLIRYAEGRILSTSA---DB 124
Db 76 YKAEVDFDRHBEIVDMIRREGS-----TRLKAEIEMGABARITLESASPARVH 126
Qy 125 GGMADSTGHKLAIVRRQVGIYLAIPYNPVNLSGSKIAPALIGNVMMKPPQGSV 184
Db 127 GIVESDVGKESRYKSAIGVGVISPMNPFILHLORSIADPALAGNAVVPASDTEV 186
Qy 185 -SGVLAKPAEAGLPAGVNTTGGSEIGDYIVHEEVPNFPGSTVQRIQGLAGMRP 243
Db 187 CGGLLARIIEEAGLPAGLFSVVGSGSEIGDAFVHPVGLVTFGTGSPVGRNIGRIAS 246
Qy 244 ---MRPIMLELGGKAGIYLAADDLNAAKOIVAGAYDSGQCTAIKRVLYVEEVD 299
Db 247 GGAHLKVALLEIGNSPFYVLGDADDEQAVNAVFGKFLHQOICAIIRIIVESDLVYA 306
Qy 358 LITPLGFHVTLDKMLAMEBPPFGPIPLIIRVYKDAEVAIAJANKSDFGQSVTRDFOKA 417
Db 366 LLAIPHVGSVTATMEIARDEIFGLRLARDEAHLALANASEYGLSAVFSRLERA 425
Qy 418 FDIANKLEVGTVHINKTGGRPNPFPLGKSGAGVQGIKRISEIEMTNKSIYLDK 465
Db 426 VAFARQLRAGMTHVNDIPVNDENAPFGGEKNSGLGRFNGDAIEBEFT 473

RESULT 16

A70318
aldehyde dehydrogenase - Aquifex aeolicus
C/Species: Aquifex aeolicus
C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

129 DASTGG--KLAIVIRROPGVILAAIPYVNVNLSGSKIAIPALIGVNVVFKPPTGSSVG 186
132 ---PGERRVGVYFRPPIGIVAITPNDPLNVAKHIGPALAAGNSLIIKESSTPLA 188
187 LVLAFAAGAGPAGVFNITITGSGEIDYIVHEBEVNFINTGSPVQRIKLAGMP 246
189 LRLVLAKEAGLPEKIVQVITGSGVIGPLTGHKRVLLISFTGQVETGKILARSAGVKK 248
247 IMELGCKADGIVLADADLNAAKQIVAGAYDYSQGRCTAIRKVLVVEEADBLAEKISE 306
249 LAMEIGSNSPTIYLQDAELMEAVAASCVSAGFCAAGNCIGVQIYVEGQSFNSFIQSFWA 308
307 NNAKLSVDPFNDAT--VTPVIDNSADFTESLVVDARQKAKELNFKDGLLTGRLD 365
309 QTKQLKLGKQSEETDIPGISKEKAKRIERVWEAKEKBARVLTGGRRTGAFETPLVT 368
366 HTYLDKLMKEEPFGILPIIRYKDAEBAVALANKSDFGLOSSVFTRDFOKAFDIANKLE 425
369 NVSPSRLLAKEAFATVILEGHSITELAIRANDVDPLQAGLFTNNILTAAPSALEKIE 428
426 VGTVINNTGGRGPNFPFLGKSGAGVQGIKYSIEAMTNVKSIVLDM 474
429 VGIWNVDSVDVRIDAMPFGIKSGLGREGVRYALIEMTBQKVAFHL 477

RESULT 19
F65045
succinate-semialdehyde dehydrogenase [NAD(P)] (EC 1.2.1.16) - *Escherichia coli* (strain K
C/Species: *Escherichia coli*
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C/Accession: F65045
R./Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of *Escherichia coli* K-12.
A/Reference number: A64720; MIMD:97426617; PMID:9278503
A/Accession: F65045
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-482 <BLAT>
A/Cross-references: UNIPROT:P25526; GB:AB000351; GB:U00096; NID:gl789011; PIDN:AACT5708.
A/Experimental source: strain K-12, substrain MG1655
C/Genetics: gabd
C/Supersfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C/Keywords: oxidoreductase

Query Match 30.3%; Score 726; DB 2; Length 482;
Best Local Similarity 34.3%; Pred. No. 1e-38;
Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;

6 QNVNNGEMKSSV--QIEILSPIDSSLGFPVPMTRBEVDHAKKAGREALPMAALTYVE 63
12 QALINGEMLDANNGEALIDVTNPANGKLSVPMGADETRAALDAANRLPAPARALTAKE 71
64 RAQYIHKAAADIERDEKEIATVLAKEISKAYNAVTEVVRTADLIRYAAEBGIRLSTSD 123
72 RATTILANFNLMHEHDDARLMTTLEQKPLAEKGEISYAASFIMWPAEBEGRIIGDPT 131
124 EGGKMDASGTHKLAIVRRQPVGIVLAIAPYVNVNLSGSKIAIPALIGVNVVFKPPTQS 183
132 PGGQAD----KRLIYIKQPIGTAITPWNFPMAAMITRAGPALAAGCTMVLKXPSQTP 186
184 VSGVLVLAFAAGAGPAGVFNITITGSGEIDYIVHEBEVNFINTGSPVQRIKLAG 243
187 FSLALAEALAIRGAVGAVFNVTGSAAGVGNLTSPNPLVKLSFTGSGEIRGLMEQCA 246
244 --NRPTMLEGCKADGIVLADADLNAAKQIVAGAYDYSQGRCTAIRKVLVVEEADBLA 301
247 KQIKVYSLGEGNAPFIVPDADLDKAVGALASKFRNAGQTCVCANRLVVDGQVYDRPA 306
302 EKISENVATLSVDPDPDN-ATVTPVIDNSADFTESLVVDARQKAKEL---NEFRDGR 357

DB 307 EKLQAVSKLHIGDGLDNVTITGPIIDEKAVAKVEHINDALEKARVVGGAHERGN 366
358 LITPLGFDHVTLDKLMKEEPFGILPIIRYKDAEBAVALANKSDFGLOSSVFTRDFOKA 417
367 FPGPTIIVDVPANAKVSKKETEGPLALPLFRFDEADVIAQANDTEGGLAAFYARDLSRV 426
418 FPIANKLEVTGTHINNTKG-RGPDNFPFLGLGSGAGVQGIKYSIEAMTNVKSIVLDM 474
427 FVGEALREYGVGIN--TGIISNEVAPFGGIRKASGLGREGSKYGLDYEITKYMIGL 482

RESULT 20
I39769
aldehyde dehydrogenase (EC 1.2.-.-) - *Bacillus stearothermophilus*
C/Species: *Bacillus stearothermophilus*
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: I39769
R./Imanaka, T.; Ohta, T.; Sakoda, H.; Wadhyastuti, N.; Matsuo, M.
J. Ferment. Bioeng. 76, 161-167, 1993
A/Title: Cloning, nucleotide sequence, and efficient expression of the gene coding for t
A/Reference number: I39769
A/Accession: I39769
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-488 <RES>
A/Cross-references: UNIPROT:P42329; GB:D13846; NID:g456298; PIDN:BA02975.1; PID:g456299
C/Supersfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C/Keywords: oxidoreductase
F/49-309/Domain: aldehyde dehydrogenase homology <ALDD>
F/255,289/Active site: Gln, Cys #status predicted

Query Match 30.2%; Score 721.5; DB 1; Length 488;
Best Local Similarity 34.4%; Pred. No. 2e-38;
Matches 166; Conservative 104; Mismatches 192; Indels 21; Gaps 9;

3 KEYQNVNNGEMKSSV-NQIE-ILSPIDSSLGFPVPMTRBEVDHAKKAGREALPMAAL 59
8 KTYFVYINGNWVSSVSNVPEPSINPANRHDIVGYQRTLEDVNEAVTANBARQTSWWR 67
60 TYVERAQYIHKAAADIERDEKEIATVLAKEISKAYNAVTEVVRTADLIRYAAEBGIRLS 119
68 SEVENGEYLYKKAHLIBGCLDIAETMRMGKTLAEAKETMRGHILIRYAGBAR-- 125
120 TSADEGCKMDASTGKLAIVIRROPGVILAAIPYVNVNLSGSKIAIPALIGVNVVFKP 179
126 ---KIGDVIIPSSDSGLLFTTRVPLGVGVISPMNFPVALPIWMAAPALVYGTVVLKPA 182
180 TCGSGVGLVLAFAAGAGPAGVFNITITGSGEIDYIVHEBEVNFINTGSPVQRIK 239
183 SETAAVAAKVIIECFHAGFPKGVNVNVCSSGVVQGIANNHPDIDGVTFGTSNTVKQVG 242
240 KLAAGR--PIMELGCKADGIVLADADLNAAKQIVAGAYDYSQGRCTAIRKVLVVEEVA 297
243 RAAPRGAKYQLEMGGKKNVIAKADADLDLAEVITSGLBSTGQKCTATSFTIEREY 302
298 DELAKISENVAKLSVDPFNDAT--VTPVIDNSADFTESLVVDARQKAKELNFKR-- 354
303 BEPKAKLIERVQKLKINGLDAETWNGPCASSQFHTVLVSYIEKGSBAGKLIYGNRCL 362
355 -----DGRLLTGLFDHVTLDKLMKEEPFGILPIIRYKDAEBAVALANKSDFGQSSV 409
363 EGBELNNGFVEEPTIEDVLDQMTIAREEIPGVALLIOVDSIEEAIKLANDREYGLSASI 422
410 FTRDFOKAFDIANKLEVGTVHINNTKG-RGPDNFPFLGLK--GSGAGVQGIKYSIEAMTNV 467
423 YTKIGNALIEFKQIDIEAGLIKNAETAGVERPQAPGCKAQSSSHSEGG--QAALIEFTSI 481
468 KSI 470
482 KTV 484

Qy	301	AKESSENVAKLSVGDPEPNAT-VTPVLDNDSADPFIESTLVYARQKAKELNEFKRDRLL	359
		302	VEVEREKVSQLOTPNAPAPATLIPGLINYOIGRIQOOLVKESSVAQKAVLEGHVOGULM
Db	360	TRGFEDHYTLMKLAMEEPGPIILPIRVKAAEVAIALNKSDDGLQSSVTPRPOKAFD	419
Qy	362	SPFLISEVTNMPAKERIFGPRIAPILKAKOBAEIAIALNDSPYGLSGSIFPGSLHRRVQ	421
Db	420	IANTLCVGTVINNKTRGGRPNFPLGLKSGGAGQGRISYIEMANTNKSIVLDMK	475
Qy	422	VAKIDITMHIHVNOQPVNAEHAISFGGKDSGIGFGEAWLIDKFTTYKWLISDEK	477

C:\Program Files\NCBI\BLAST\bin

```
leach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.  
Koesterer, P.; Koningsstein, G.; Krovcin, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A.; Authors: Lauber, U.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Mauel  
Y., M.; Ogawa, K.; Ogiwara, A.; Outega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetlie  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,  
M.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Seguyuchi, J.; Sekowska, A.; Seron  
Aouchi, M.; Tanakoshi, A.; Tanaka, T.; Terpetra, P.; Tognooni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, K.  
A;Authors: Yoshikawa, H.F.; Zunnstein, E.; Yoshikawa, H.; Danchin, A.  
Article: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
Accession number: A69580; MUID:98044033; PMID:9384377  
A.Accession: C69584  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A.Molecule type: DNA  
A.Residues: 1-485 <KUN>  
A.Cross-references: UNIPROT:P94358; GB:Z59123; GB:AL009126; NID:g2636240; PIDN:CAB15909.  
A.Experimental source: strain 168  
C.Genetics:  
C.Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
```

C:\Program Files\NCBI\BLAST\bin

```
Query Match          28.1%; Score 673; DB 2; Length 485;  
Best Local Similarity 33.9%; Pred. No. 2.4e-35;  
Matches 161; Conservative 95; Mismatches 201; Indels 18; Gaps 6;
```

C:\Program Files\NCBI\BLAST\bin

```
OY QNYNGEWM--SSVNOIEILSPIDDSSLGFVPAMTREEDHAMKAGREALPMAALTYYE 63  
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB   8 KSFINCKMTGSGESGRTEIDLNPQDSVYTASLATKGLEDPAFIAQAOKEMAXSTED 67  
  
OY 64 RAQYLHKADIIRDKKEILTAVLAKETSKAVNSYTEVVRA---DLIRVAEGSIRLS 119  
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB   68 RAAVLLOKARGCYLHENRDDITIMTAIRETGTTIKSTEILEQTIALDEAMTYIGELGVYE 127  
  
OY 120 TSADEGGKMDSATGHKLAVIROPGVIPLAAPNVYPNLGSKIAPALLIGNVMVFKP 179  
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB   128 VSBDISGKTNK-----IYRLDPLGYISSISPFPFNMLSRSLAPAILANSVVHKPD 179  
  
OY 180 TGQSVS-GLVLAFAEAAGLPAGVFNTITRGSEIGDYIVHEBEVNFIFTGSTPVGRI 238  
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB   180 IQTAISGGTIATAFEHAGLPAGVLNWLTDVKEIGDGLTNPIRLISTGSTAVGRHI 239  
  
OY 239 GKLAG-MRPIMELEGKDGIYLDADLDNAAKOIYAGAYDSQRCTAIRVLVEEV 296  
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB   240 GETIAGAFAFKMALDELGNPNPFAVLSDDVDRAVAAIKGKFHQQCIMIINRIITHOV 299  
  
OY 297 ADIELAKISEENVAKLSVGDPFDNATVT-PVIDDNSADPTIESLVVAROGKAELNEPKRD 355  
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB   300 YDFEFVEKTKARKVKOLPYGDOTDPTKVGPFLNEROTEALEITEBAKTDGIELAVEGRV 359  
  
OY 356 GRLLTFQLPDHWLDMKLAWEEPPGPIILPIRVDAEBAVALANKSDFOLOSSVFRDPO 415  
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB   360 GNVLTLTVYFGADNNKSIIAQTELPAPIATIINKASDGDAIDMANTEVGLSAVFTSDE 419  
  
OY 416 KAFDIANKLEEVGVTHINNKTGRGPDNPFPLDKSSGAGVOGIRYSIEANTVXSI 470  
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB   420 KGEEKPALQIDSGMTHVNDOSVNSDNIAFGCNKASGVGRFGNPMVVEBFTWKMI 474
```

C:\Program Files\NCBI\BLAST\bin

```
RESULT 34  
A87547  
vanillin dehydrogenase [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004  
C:Accession: A87547  
R.Niemann, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.U.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolton  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Ventner, J.C.; Frazer, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A>Title: Complete Genome Sequence of Caulobacter crescentus.  
A.Reference number: A87249; MUID:21173698; PMID:11259647  
A.Accession: A87547  
A.Status: preliminary  
A:Molecule type: DNA
```

A:Residues: 1-470 <STO>
A:Cross-references: UNIPROT:Q9A5F5; GB:AE005673; NID:gl3423939; PIDN:AAK24373.1.; GSPDB:CC2402
A:Gene: CC2402
A:Genetics:
C:superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 28.0%; Score 669.5; DB 2; Length 470;
Best Local Similarity 34.9%; Pred. No. 3: 5e-35;
Matches 155; Conservative 88; Mismatches 190; Indels 11; Gaps 4;

QY 36 AMTREVVDHAMAGGEALPAMAAALTVYERAOYLHKADIIERDKEIATVLAKEISKAYN 95
DB 24 AMTADQARAIVDAQAALPTWSALGNARALLSKAAALAEARDDFVAAMWGEIGATEG 83
QY 96 ASVTEVETADLIIRAAEBGIRLSTADSGGKMDASTGHKLAVIRROPGVILAIAPYNY 155
DB 84 WAFPMILMAASVYREAA-----ALTTOVSGEVIPSIDPGCIAMAVREBVGILGIAPWNA 138
QY 156 PUNISGSKIAPALIGGNVMPKRPPTQSVSGVLAKFAEAGLPAVGNITGR--GSE 212
DB 139 PILLVRAAVATPLACNTIVLAKASECPRTHEILAEFAAAGLEBALSTVYNAPADAE 198
QY 213 IGDYIVHEEVEVNFINEFTSGTPVQIRIGKLAG--MRPIMELEGKAGIVLADADLDNNAK 270
DB 199 VVGLALIDHPAAVRINFTSGTAVGKI IAKRAAEHLKPVLLLEGKAPLIVLEDADDEAVK 258
QY 271 QIVAGAVDYSGRCRTAIKVLVVEEVADELAEKI SENVAKLSVGDPEPDNAF-VTPYIDN 329
DB 259 AAFAFAMVQGI CMSTERIIIVDAIADAFVAKFAAKASSIAGVDPEGKTPLAGVVLK 318
QY 330 SADFIESIIVDAROGAKELNEFRDGRLLTPGLFDHVTLDMKAWPEPFGIPIIRVK 389
DB 319 TVTHVOGLAADLADCAVGVSGGPANGVLMPTAVDVYTPAMKLFREESFEPVAIVAK 378
QY 390 DAEBAVALANKSDFGLSSVFTRDPQAPDIANKLEVGVTHINNKTRGKPDNFPFLGLKG 449
DB 379 DEEHAIALANDREYGLSASVFTRDIARGLKVARIGICIVNPGPTYVDAEAMPGGVKA 438
QY 450 SGAGVQGI RYSTEAMTVNKSIVLD 473
DB 439 SGYGRFGKAGVDAFTELRMITVE 462

RESULT 35
DB3613 succinate-semialdehyde dehydrogenase PA0265 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: DB3613
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
atman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Lm,
ature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: DB3613
A:Status: preliminary
A:Molecule type: DNA
A:Reads: 1-483 <STO>
A:Cross-references: UNIPROT:Q916M5; GB:AE004464; GB:AE004091; NID:g9946099; PIDN:AAG0365
A:Experimental source: strain PA01
C:Genetics:
A:Gene: gabD, PA0265
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 27.9%; Score 668; DB 2; Length 483;
Best Local Similarity 33.2%; Pred. No. 5e-35;
Matches 158; Conservative 90; Mismatches 212; Indels 16; Gaps 7;

QY 6 QNYNNGEKSSYV--QIEILSPIDDSSLGFPVPMTRREVVDHAMAKGREALPAMALTYE 63
DB 12 QNYVGAWVDANQGQTIKVNNPPTGIIIGSVPKGAETRAIRLPAADLALPMARLITAKE 71

```

QY 64 RAQVHLHKAADIIERRDKESIALATVLAKEISAKANASTEVVRADLIRVAEEGILNLS1SD 123
Db 72 RANKLRPFMDLMTENQDDLARLMTIEQCKPLAEAGELVAAYASFLEWGEBAKRIYGD7I 131
QY 124 EGGKMDASTGHKLAVIRQPVGIVLAIAPYVNVLSGSKTAPALIGNVVMPFPPTGGS 183
Db 132 PGHQPD-----KRIIVIKQIPIGVTRAITPMPFPSMIRKAGPINALAGCTVLRKASGTP 186
QY 184 VSGVLAKAFABAGLPACVFNPTTGRGSEIGDYIYEHEEVNFNFTGSTPVGORIGK--L 241
Db 187 YSALATAEIABERAGIPKGVFSVVTGSGAGEVGBELTSNPIYAKLFTFGSTETIGRQIMABCA 246
QY 242 AGMRPIMELGCKDAGIVLADADLDNNAKOIVAGAYDYSGORCTRAKRVLYVEEYABELA 301
Db 247 QDIKVSEIETLGGNAPFIYFDDADLDAAVEGALISKTRNNQDTCVCANELLVYQDGYDAFV 306
QY 302 EKISENVAKLGVGDFFD--NATVTPVIDDNSADFIESLTVDDARKGAKXELNEFK---RDGR 357
Db 307 DKLRKAAYAKKLINGLEAGVTTGPLIDAKAAKAYEBEHADAVSKAKVVSQKPRHALGCT 366
QY 358 LITPGLFDVHTLDMKLAMEBEPGPILPIIRVDAEBAVALANKSDPGLQSSVFTDFEOKA 417
Db 367 FFEPTIIVDVDPKNALVSKDETFGPIAPVFRFRDEBAFVLAMSNDRIFGLASIFVARDLARV 426
QY 418 FDIANKLEVGVIHINNTGR--GPDNFPFLGLKGSAGVQGIIRYSIEAMTNVKSIVL 472
Db 427 FRVIEQLEIYGMWGIN--TGLISNEVAPFGGIKASGLGREGSKYGIEDLETIKYCL 480

```

```

RESULT 36
E75574
aldehyde dehydrogenase - Deinococcus radiodurans (strain R1)
C|Species: Deinococcus radiodurans
C|Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C|Accession: E75574
R|White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Dodson, J.D.; Dodeon, R.J.;
S.; Smith, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Peterback, T.; Zalewski, C.; Mc-
Science 286, 1571-1577, 1999
A|Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A|Reference number: A75250; MUID:20036896; PMID:10567266
A|Accession: E75574
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-524 <WHI>
A|Cross-reference: UNIPROT:O9RYT8; GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF1241
A|Experimental source: strain R1
A|Genetics:
A|Gene: DRA0220
A|Map position: 2
C|Superfamily: NAD-dependent aldehyde dehydrogenase, aldehyde dehydrogenase homology

Query Match 27.9%; Score 666.5; DB 2; Length 524;
Best local Similarity 34.4%; Pred. No. 6.9e-35;
Matches 166; Conservative 79; Mismatches 222; Indels 15; Gaps 7;

QY 1 LTKRYQNVNGEW--KSSVNOIETLSPIDDSSLGTFVPMTRREEVHAMKAGREALPAMAA 58
DB 27 LSRGIQHRTIGGEWVVAHSGKTPDASHPVNDNLVVAAGDASDIRAKAAHDAFQYRE 86
QY 59 LTVERRAOLYLRKADIIRDKKEIATVLAKEISKAYNVASTEVTADLIRYAAEGIRL 118
DB 87 VSGARRRKLHLKADLIEKAEIIVLESVDTGQAIRFKKSAARGAENFRYADR---- 142
QY 119 STSADGGKMDASTGHKLAIVTRQPVGIYLAIAFNRYPNVNSGSKIAALLIGANVMYKP 178
DB 143 APGADGGSLRPPGFINSI--RQPIGPVGVITPNNTPFMTSTWKIAPALAAAGCTVAKP 200
QY 179 PTQGSVSGVLAKAFEAAGLPAGVFNTITGSGEISGDIYVEHEEVNFINTGSPVQRI 238
DB 201 AEMSVASATLLEIMDEAGLPBGVNNLVHGFSGSKSLTEHPLYKVAAPVGETTGSHT 260
QY 239 GKLAG--MRPIMLELGGKAGIVLADADLNNAAKOIVAGAYDYSGORCTAIRVLVEEV 296

```

Db	261	MRGADTLTKRVHVELOGKNPVVFPDDADLDKLDLAVFMVYSLNGERKCTSSRVLIDBGI	320
QY	297	ADELAERKISENVAKUSVEVDPEFNAT-VVPVYIDONSADFIESTLVYDARQKGAKEINERK-	354
Db	321	YDFETRIIRERANRIRVGDPLDPDPTREVBGLVHPRHFEKVMSEYFECARQEGATTAAQGERV	380
QY	355	--DGRLLTQGLFDHVTLLDMKLAEBEPPFPIIIRVKDAEBAVNAKSPQGLQSSVETR	412
Db	381	GEAGNFVRPTLFTAAANDMKIAQEEIFGCVTLAIPFKDEALALBLADVNGVLAGLAVMTN	440
QY	413	DPQAKFDINKLVEGVVHINNTKGGSDPNPFLIGKSGCAVQGIIRYSIEAMTNVKSIVL	472
Db	441	DLRAIRFAHGLEAGMIWVNSENVHLP-PPGVKNSGIGRDGDYSFFETMETKNIAI	499
QY	473	DM	474
Db	500	SL	501

[illegible]

C:Genetics:
A:Gene: betB
A:Map position: 7.5 min
C:Complex: homotetramer
C:Function:
A:Description: catalyzes the oxidation of betaine aldehyde to betaine using NAD+ and water
A:Pathway: betaine biosynthesis
A:Note: betaine is a protective osmolyte induced to accumulate by osmotic stress
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: alternative initiators; homotetramer; NAD; oxidoreductase; stress-induced pr
F:2-490/Product: betaine-aldehyde dehydrogenase #status experimental <MAY>
F:45-306/Domain: aldehyde dehydrogenase homology <ALDD>
F:252,286/Active site: Glu, Cys #status predicted

Query Match 27.5%; Score 658.5; DB 1; Length 490;
Best Local Similarity 31.9%; Pred. No. 2e-34; Indels 35; Gaps 9;
Matches 157; Conservative 101; Mismatches 199;

QY 4 EYONTYNGEMKSVN--QIEILSPIDDSIGVPVPMTRREVDHAKAGREALPAMALTY 61
DB 6 EQQLYHGGYTATSGRTFETINPANGVNLATVQAAGRDVDRAVSAQGGKIMASMTA 65
QY 62 YERQYLAKAADIIERDKKEIATVLAKEISKAYN-ASVTEVVRTADLIRYAAEGIRLST 120
DB 66 MERSRIIRRAVDILBERNDLAKLETLDGKAYSETSTVDIVTGADVLEYTA----- 117
QY 121 SADEGGKMDASTGHTAV-----IRROPVGIATLAIPYVYNLSSGKIAPALIGNV 173
DB 118 -----GLIPLBGSQIPLAETSFYTRRSPFLGVAGIGMNPDIQIALMKSPALAAAGNA 172
QY 174 VMFKPTPGSSVGLVTLAKAPAEAGLPAGVFNITIGRSBIDGYIYEHEVNFINTGSRP 233
DB 173 MIFKSEVPLPLALKLAIEYSEAGLPDGFVNLPGVAGTGGQLTEHPGIKAVSFTGVA 232
QY 234 VQQR-----GKLAGMPIMLELGKAGIYTLADADLDDNAKQIVAGAYVYSGORCTAIRKY 290
DB 233 SGKTKWANSAASSLKRVETMELGKSPLIYFDADLDLADIMMANFFSGGVCVCTGTGV 292
QY 291 LVVEEVADELAKISEBNVAKLSVGDPFDVAT--VTPIIDNSADFIESTLVADAROKGAKEL 349
DB 293 FVPACKAAFEOKILARVIRIAGVDFDQTNFGLVSPPHNDNVLRYAKGEBGARVL 352
QY 350 ---NEFKRD---GRLLTPGLFDHYTLDMKLAMEBEPFGIILPIIRYKDAEAAVALANKSD 402
DB 353 CGGDLVKGDPNGAMVAATFTVDCDDMTIVREELFGVMSILTVESDEVIRRANDTD 412
QY 403 FGLQSSVFRDFOKAFDINKLEVEGVTHINNKTRGPDNPPPLGLKGSAGVQGIIRYSIE 462
DB 413 YGLAAGIVADLNRRHVRVHOLEAGICWINT--WGESPAPMPVGYGHSIGIRENGVMTLQ 471
QY 463 AMTNVKSITLDM 474
DB 472 SITQVKSIOVEN 483

RESULT 45
DEHUB1
aldehyde dehydrogenase (NAD) (EC 1.2.1.3) 1, cytosolic [validated] - human
N:Alternate names: aldehyde dehydrogenase E1
C:Species: Homo sapiens (man)
C:Date: 28-May-1986 #sequence revision 31-Dec-1993 #text change 09-Jul-2004
C:Accession: A33371; I51846; I39431; JCS554; A00363; A61010; A26817
R:Hsu, L.C.; Chang, W.C.; Yoshida, A.
Genomics 5, 857-865, 1989
A:Title: Genomic structure of the human cytosolic aldehyde dehydrogenase gene.
A:Reference number: A33371; MUID:90077427; PMID:2591967
A:Accession: A33371
A:Molecule type: DNA
A:Residues: 1-501 <HSD>
A:Cross-references: UNIPROT:P00352; GB:M31594; NID:gl78370; PIDN:AAA51692.1; PID:gl78372
R:Yoshida, A.; Hsu, L.C.; Yanagawa, Y.
Adv. Exp. Med. Biol. 328, 37-44, 1993
A:Title: Biological role of human cytosolic aldehyde dehydrogenase 1: hormonal response,

A:Reference number: I51846; MUID:93263033; PMID:8493914
A:Accession: I51846
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6 <YOS1>
A:Cross-references: GB:S61235; NID:G300399; PIDN:ADI3925.1; PID:G4261625
R:Yoshida, A.; Ikawa, M.; Hsu, L.C.; Tanl, K.
Alcohol 2, 103-106, 1985
A:Title: Molecular abnormality and cDNA cloning of human aldehyde dehydrogenases.
A:Reference number: I39431; MUID:85252089; PMID:4015823
A:Accession: I39431
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1, 163-501 <YOS2>
A:Cross-references: GB:M26761; NID:gl78393; PIDN:AAA55518.1; PID:gl78394
R:Kathmann, E.C.; Lipicky, J.J.
Biochem. Biophys. Res. Commun. 236, 527-531, 1997
A:Title: Cloning of a cDNA encoding a constitutively expressed rat liver cytosolic aldehy
A:Reference number: JCS554; MUID:97382470; PMID:9240474
A:Accession: JCS554
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-120, S', 122-501 <KAT>
A:Cross-references: DDBJ:AF003341; NID:G2183298; PIDN:AACS1652.1; PID:G2183299
A:Experimental source: liver
R:Hempel, J.; von Bahr-Lindstrom, H.; Jornvall, H.
Eur. J. Biochem. 141, 21-35, 1984
A:Title: Aldehyde dehydrogenase from human liver. Primary structure of the cytoplasmic i
A:Reference number: A00363; MUID:84208019; PMID:6723659
A:Accession: A00363
A:Molecule type: protein
A:Residues: 2-501 <HEM>
R:Argawal, D.P.; Cohn, P.; Goedde, H.W.; Hempel, J.
Enzyme 42, 47-52, 1989
A:Title: Aldehyde dehydrogenase from human erythrocytes: structural relationship to the
A:Reference number: A61010; MUID:89377753; PMID:276714
A:Accession: A61010
A:Molecule type: protein
A:Residues: 23-27; 79-85; 101-107; 114-128; 132-142; 144-154; 309-319; 421-434; 477-483 <AGA>
A:Experimental source: erythrocytes
R:Abriola, D.P.; Fields, R.; Steinh, S.; Mackereit Jr., A.D.; Pietruszko, R.
Biochemistry 26, 5679-5684, 1987
A:Title: Active site of human liver aldehyde dehydrogenase.
A:Reference number: A26817; MUID:88050861; PMID:3676276
A:Accession: A26817
A:Molecule type: protein
A:Residues: 266-273 <ABR>
A:Note: the active site Glu was identified by suicide inhibition with bromoacetophenone
C:Genetics:
A:Gene: GDB:ALDH1
A:Cross-references: GDB:119667; OMIM:100640
A:Map position: 9q21-9q21
C:Complex: homotetramer
C:Function:
A:Description: catalyzes oxidation of an aldehyde to an acid using NAD+ and water
A:Note: enzymes with this activity are involved in diverse metabolic pathways in various
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: acetylated amino end; alcohol metabolism; cytosol; homotetramer; liver; NAD;
F:2-501/Product: aldehyde dehydrogenase (NAD+) 1 #status experimental <MAY>
F:59-323/Domain: aldehyde dehydrogenase homology <ALDD>
F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental
F:269,303/Active site: Glu, Cys #status experimental
F:456/Binding site: NAD (Cys) #status experimental

Query Match 27.5%; Score 658; DB 1; Length 501;
Best Local Similarity 34.1%; Pred. No. 2.3e-34;
Matches 163; Conservative 85; Mismatches 210; Indels 20; Gaps 8;

QY 8 YVNGKMSVSN--QIEILSPIDDSIGVPVPMTRREVDHAKAGREAL---PAMALTY 62
DB 24 PINNMHDSVSGKFPVNPATREELCOVEEGDKEDVAKAAYAAQAIQIGSPWRTMDAS 83
QY 63 ERAQYLAKAADIIERDKKEIATVLAKEISKAY-NASVTEVVRTADLIRYAAEGIRLST 121

```
Db      84 EKRRLLYKLAADLERERRLALATMESNMGKLYNSAYLNDLACITKILRYCAGMADKIQ-- 141
Qy      122 ADEGKMDASTGHKLAVIRROPGIVLALAPYVPVNLSSGSKTAPALLIGNVVMFKRPPTQ 181
Db      142 ----GRTIPIDGNPFYTRHEPIGVCGQIIPNNFPLMLKIKGPAISGNTVVAPEAQ 197
Qy      182 GSVSGVLAKAFPAEAGIPAGVNTITGRGSEIDYIVHEEVNFINTGSTPVGQRIKGL 241
Db      198 TPLTALHVASLIEAGPPGVNIVPGYPTAGAAISSHMDIDKVAFTGSTEVGKLIKEX 257
Qy      242 AG---NRPIMLEEGKAGIVLADADLDNAAKOIVAGAYDSGORCTAIRVYVEEVAD 298
Db      258 AGSNKRYLTELGGSPCTVLADADDNNAVEFAHHGVFHYGQCCCLASRTVEESITD 317
Qy      299 ELAEKISENVAKLSVSDPF-DNATVTPVIDNSADFIESTLVVAROKGAK---ELNEFKR 354
Db      318 EFRRSVERAKKYILNPLTPGVTOGPQIDKEQYDIKLIDIESGKKEGALTECGGPGW 377
Qy      355 DGRLLTPGLFDHYTLDMKLAMEEPPGPIILPIRVKQAEBAVAIAKNSDFGLQSSVFTTRDF 414
Db      378 KGYVQPVFVSNTDEKRIAKEEIFGVQOIMKFKSLDDVIRKRNNTFYGLSAGVFTKDI 437
Qy      415 OKAFDIANKLEVGTVHNNKTGRGPDNFPPLGLKSGAGVQGIIRYSIEAMTNKSIYL 472
Db      438 DKAITISSALQAGTAVW-NCYGVVSAQCPFGFGRMSGNGRELGEYGFHEYTEVKTIV 494
```

RESULT 46

```
C95964
probable aldehyde dehydrogenase protein (EC 1.2.1.-) [imported] - Sinorhizobium meliloti
C/Spectes: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: C95964
R/Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chalin, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A/Reference number: A95842; M0ID:21396508; PMID:11481431
A/Accession: C95964
A/Status: preliminary
A/Molecular type: DNA
A/Residues: 1-485 <KUR>
A/Cross-references: UNIPROT:Q92UV7; GB:AL591985; PIDN:CAC49379.1; PID:G15140865; GSPDB:C
A/Experimental source: strain 1021, megaplasmid pSymB
R/Galibier, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chalin, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A/Reference number: A96039; M0ID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics:
A/Genes: SMD21539;
A/Genome: plasmid
C/Suprafamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C/Keywords: oxidoreductase
```

```
Query Match      27.4%; Score 656; DB 2; Length 485;
Best Local Similarity 36.1%; Pred. No. 2,9e-34;
Matches 171; Conservative 84; Mismatches 201; Indels 18; Gaps 8;
Qy      9 VNGEMKSVNQIILSPIDSSLGFPVAMTREEDVHAMKGRALPAMALTYTERAOYL 68
Db      17 IADRVLDTDRVVRKPMNDTVGVTPA---GRAEHRAREFAIAAYOPGLTYTEROKIL 73
Qy      69 HKAADIIERDKKEIATVLAKEISKAVNASVTEVTRADILRYAAEGIRLSTADEGGRK 128
Db      74 LATAEALAAKEEISIDVITLELIGSKADSLYEVGRAFDVFTLAGQMCIR-----DDGEIF 128
Qy      129 DAS-TGHKLA---VIRROPYGLATAPYVPVNLSSKTAAPALLIGNVVMFKRPPTQGSV 184
```

```
Db      129 SCDLTFPHGARKIFTRREPLTAISAITPPNHNLANVAHVAALATNNCVVVFTELTPM 188
Qy      185 SGLVLAKAPAEAGLPAGVNTITGRGSEIDYIVHEEVNFINTGSTPVGQRIKGLAGM 244
Db      189 TALLADLIYEAGLPPEMLSVTGWPRADIGEMBITPHVDLVTFTGSVPVGKLIANAHY 248
Qy      245 RPIMLELGGKDAQIV---LADADLNAAKOIVAGAYDSGORCTAIRVYVEEVADLA 301
Db      249 KQVLELGGNDPLILNDLSDDLAARADLAAGATKNSGQCTAVAKRILCOESVADRPV 308
Qy      302 EKISENVAKLSVSDPFDNAT-VTPVIDNSADFIESTLVVAROKAGEINERDRRLTT 360
Db      309 PLVTERAKRLRGDPRDRDITDGTIVHEKAAALFEERVWRABEGADILYHPOSSALLP 368
Qy      361 PGLFDHYTLDMKLAMEEPPGPIILPIRVKQAEBA-VALANKSDPGLQSSVFTTRDFQKAD 419
Db      369 PLYVDRVPHQSDLVLEETFGPIILPIRVKQAEBA-VALANKSDPGLQSSVFTTRDFQKAD 428
Qy      420 IANKLEVGTVHNNKTGRGPDNFPPLGLKSGAGV-QGIRYSIEAMTNKSIYL 472
Db      429 YIAGLKVGTVNIEWPVGRIEMSFGGIRDSGNGYKEGVIEAMKSFYNKTESL 482
```

RESULT 47

```
AD3300
sucinate-semialdehyde dehydrogenase [NAD(P)] (EC 1.2.1.16) [imported] - Brucella melitensis
C/Spectes: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C/Accession: AD3300
R/Delvecchio, V.G.; Kapatal, V.; Redkar, R.U.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldman, E.; Selkov, E.; Elser, P.H.; Hagius, S.; O'Callaghan, D.; Letes-
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen Brucella meliten-
A/Reference number: AD3252; PMID:11756688
A/Accession: AD3300
A/Status: preliminary
A/Molecular type: DNA
A/Residues: 1-487 <KUR>
A/Cross-references: UNIPROT:Q8YI07; GB:AE008917; PIDN:AAL51567.1; PID:G17982287; GSPDB:A
A/Experimental source: strain 16M
C/Genetics:
A/Genes: BMEI0386
A/Map position: 1
C/Suprafamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C/Keywords: oxidoreductase
```

```
Query Match      27.4%; Score 655; DB 2; Length 487;
Best Local Similarity 35.3%; Pred. No. 3.4e-34;
Matches 169; Conservative 81; Mismatches 201; Indels 28; Gaps 11;
Qy      6 QNVVNGEMKSVN--QIEILSPIDSSLGFPVAMTREEDVHAMKGRALPAMALTYVE 63
Db      13 QCLVNGRMWDADADGTTIKTNPADSGVIGTVELSATIKKALIDSAKALSGMAATKARE 72
Qy      64 RAQYTHKAADIERDKKEIATVLAKEISKAVNASVTEVTRADILRYAAEGIRLSTAD 123
Db      73 RAGILRKMFDDLITANADIALIMTSEOGKPLAEARGEVLVYASFTEWFAEAKRV----- 127
Qy      124 EGGKMDA-STGHKLAVIRROPGIVLALAPYVPVNLSSGSKTAPALLIGNVVMFKRPPTQ 182
Db      128 YEDTIPAPONGRLTVI-RQPVGVDAITPMMPAMITRKAAPALAACTMIVRADLT 186
Qy      183 SVSGVLAKAFPAEAGLPAGVNTITGRGSEIDYIVHEEVNFINTGSTPVGQRIKGLA 242
Db      187 PLTALALGILAEKAGIPAGVLOIVGKAREISABLSNTVAKLSTGSTVGRLL--WA 244
Qy      243 GRP-----TMELEGKDAQIVLADADLNAAKOIVAGAYDSGORCTAIRVYVEEVAD 298
Db      245 OCAPTIKRISLIEGNAAPFIVPDADLDAAVGAWSKTRNAGQTCVCANRTYVORGVD 304
Qy      299 ELAEKISENVAKLSVSDPDF-NATVTPVIDNSADFIESTLVVAROKAGEINERDRDR 357
Db      305 KFAEKILAAVYKELKVGNGTBPVGVIIGPMIEEKAITVKAHIEDAVSKAKLITGGEKEGG 364
```


R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatawara, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; X
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A:Reference number: A72450; PMID:9310339; PMID:10382966
A:Accession: H72562
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-485 <RAW>
A:Cross-references: UNIPROT:Q9YB09; DDBJ:AP000062; NID:G5105244; PIND:BAAB0789.1; PID:95
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1786
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 27.2%; Score 651.5; DB 2; Length 485;
Best Local Similarity 33.9%; Pred. No. 5.6e-34;
Matches 163; Conservative 93; Mismatches 210; Indels 15; Gaps 7;

QY 5 YONVYVNGEWSVNOQIEILSPIDDSLGIV----PAMTREVDMHAKAGREALPAMAAIT 60
DB 2 FTFVAGNVNVTLSNLAIVVRSPIDDSIIANKSRDLAAALEBGVDVYRIGRRRI--RDVP 58
QY 61 VYBRADYVHKADIIERDKKEIATVLAKEISKAVNASVTEVTRADLIRYAAEGIRLST 120
DB 59 GEKRLNIIILRIADILEKNADDFEEVLITVNAKTRKQAAAGEVASIDRLKALD-LRKIQ 117
QY 121 SADDEGKNMDSGTHKLAIVTRQPGVITATAFYNYPNVLSGSKTAPALIGGVNVPKPT 180
DB 118 GEYVPDQWRHRTLESSEGIYRREPVGVALIIPFNYPLEDTYNNKFVYSTIPGAAFTKPS 177
QY 181 QGSVSGLVAKAPAEAGLPAGVFETITGRGSEIDYIEHEEVNPFNTGSTPVGORIGK 240
DB 178 ADPLPYIMARALALEAGFPPEESLMATLRGSE-AEKLVADRIRISITLSGSETGVKWR 236
QY 241 IAGMRPIMLELGSKDAGIVLADADLDNAAKQIVAGAYDYGQORTAIRKVIIVEEVADEL 300
DB 237 SAGIKQILMELGGSDPAIVLADADLKSAASKIATGITSYGQRCDAITLVIVEEPVNM 296
QY 301 AEKISNVAKLSTGDP-FDNAATYTPVIDNDSADTIESLVNARQGAELNEFKDG-RL 358
DB 297 KKLVLVELSKIKVGGDPNDSTITMGPLIDKEAVDAMVEAIVIEBASVGAQLLVGGERIGDTY 356
QY 359 LTPGFLD---HYTLDMKLAMEEPFGPIPLPIRYKDAEVAVALANKSPGQSSVETBDF 414
DB 357 TTPALVEVEDKVALREIKLYKDEYFAFVAVITSTLDEALIELANGRRYGLDAALFGESM 416
QY 415 QKAFDIANKLEVGTVHINNTKGRGPDNFPFLGLKSGAGVQIGIRYSIEAMTNVKSIVLDM 474
DB 417 AIRIKLIRFLEVGAITYNEYPRHGIGYYPFGGRDGSIGRRGIGVSIYVALAKTIVNY 476

QY 475 K 475
DB 477 R 477

RESULT 51
G75592
succinate-semialdehyde dehydrogenase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: G75592
R:White, O.; Egan, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; PMID:20036896; PMID:10567266
A:Accession: G75592
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-487 <WHI>
A:Cross-references: UNIPROT:Q9RZB6; GB:AE001862; GB:AE001825; NID:G6460468; PIND:AAF1229

[illegible]

C:Genetics:
A:Gene: gabD4
A:Genome: plasmid
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 27.2%; Score 650; DB 2; Length 490;

Best Local Similarity 33.6%; Pred. No. 7.1e-34;
Matches 160; Conservative 84; Mismatches 214; Indels 16; Gaps 9;

```
QY 9 VNGEM---KSSVNOGIELSPIDSSLGFPVPMTEBEVDHAKAGREALPAAALTYERA 65
DB 23 IGAEMLDNRDSGKTPVSNPATGEVYALIPDMSRSETARAIIDAAHAAQMAEKTKERA 82
QY 66 QYLAHRAADIIEERDKKEIATVLAKEISKAYNASVTEVTRADLRVAEEGIRLSTSDRG 125
DB 83 AVLRLVLDLVANADDDIITLMEKPKLTEAKGELLYGASVYEMFGEAKVYGGDTTIG 142
QY 126 GMDASTGHKLAIVRRQPVGIYLAIPYNYPVNLSSGKIAPALLIGNVMPKPPQGSYS 185
DB 143 HQPD-----KRITIVKQPIGVVAALTPMNFPMMLARKLAPAAAGCAVSKPAETPLS 197
QY 186 GLVLAKEAFEAELPAGVFNTITGRGS-ELGDIYVEHEBNVNFNTGSGTSGVIGRLAGM 244
DB 198 ALALALLAERAGLPAGVFENVILSTDSAEVKGEMCANDKVRKLTFTGSTIVG-KILMROGA 256
QY 245 RPIM---LELGSKDAGIVLADADLDNAKQIVAGAYDYGQRCTAIKRVLYVEEVADELA 301
DB 257 DQIMKLGLELGNAPFIVDDADLDAAVAGAAVAKTRNNQTCVCANRIFVQAGIYDAPA 316
QY 302 EKISENVAKLSVGDPPD-NATVTPVIDNSADFIESLVVDARQAK-ELNEFKDGRLL 359
DB 317 ARLTAKVSEMTIGDGEPPVDAGPLISEKALAKVEEHIRDAVTYKADLVLGNAAGLFF 376
QY 360 TEGLDHYTLDMKLAMEEPFGPLPIIRKDAEBAVALNKSDPGIOSSVFTRDQKAD 419
DB 377 EPTVLTGATMDKIKAGEETFGVAPLFEKETEENEVSMANKTEFGIASYFSKDVSKYFR 436
QY 420 IANKLEVGTVHINNKTGR-GPDNFPPLGKSGAGVQGIIRYSIEAMTANKSVLDM 474
DB 437 VAEALFEGMVGIN--TGLISTEVAFPGVKQSGQREGSKTIGDYEYTKYCLSI 490
```

RESULT 53

A55684
aldehyde dehydrogenase (NAD) (EC 1.2.1.3) 6 precursor, salivary - human
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1995 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C:Accession: A55684
R:Hsu, L.C.; Chang, W.C.; Hiraoka, L.; Hsieh, C.L.
Genomics 24, 333-341, 1994
A>Title: Molecular cloning, genomic organization, and chromosomal localization of an ad
A:Reference number: A55684; MUID:95213025; PMID:7698756
A:Accession: A55684
A:Molecule type: mRNA; DNA
A:Residues: 1-512 <Hsu>
A:Cross-references: UNIPROT:P47895; GB:U07919; NID:9995897; PIDN:AA79036.1; PID:9544482
C:Comment: This isozyme is found at highest levels in saliva, stomach, and kidney and at
C:Genetics:
A:Gene: GDB:ALDH6
A:Cross-references: GDB:364103; OMIM:600463
A:Map position: 15q26-15q26
A:Introns: 33/3; 66/3; 115/3; 159/1; 179/3; 222/3; 260/3; 295/1; 356/3; 411/3; 464/2; 48
C:Function:
A:Description: catalyzes oxidation of an aldehyde to an acid using NAD+ and water
A>Note: enzymes with this activity are involved in diverse metabolic pathways in various
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: NAD; oxidoreductase; saliva
F:70-334/Domin: aldehyde dehydrogenase homology <ALDH>
F:280.314/Active site: Glu, Cys #status predicted
F:467/Binding site: NAD (Cys) #status predicted

Query Match 27.2%; Score 650; DB 1; Length 512;

Best Local Similarity 34.0%; Pred. No. 7.5e-34;
Matches 164; Conservative 85; Mismatches 208; Indels 26; Gaps 9;

```
QY 8 YVNGEWSNV--QELISPIDSSLGFPVPMTEBEVDHAKAREL---PAAALTY 62
DB 35 FINNEMHBSKSGKFPATCPSTRBOICEVEBGDPDVKAVALQVAFQSGSPMRRLAL 94
QY 63 ERAQYLAHRAADIIEERDKKEIATVLAKEISK---AYNASVTEVTRADLRVAEEGIRL 118
DB 95 SRGRLLHDLADVERDRATLALETMDGKFFLAFFIDLEGICRT---LAYPMAMAKI 151
QY 119 STSADGSKMADSTGHKLAIVRRQPVGIYLAIPYNYPVNLSSGKIAPALLIGNVMEKP 178
DB 152 Q-----GKTIPTDNNVCFTRHEPIGVCGAITPMNPFLLMLWMKLALCCGTMVLKP 205
QY 179 PTQGSVGLVLAKAPAEAGLPAGVFNTITGSGELGDIYVEHEBNVNFNTGSGTSGVIGRI 238
DB 206 ABQTPVLTALVIGSLIKGFPFGVNVIVPGFPTVGAALISSHPQINKLAFTGSTEVEGLV 265
QY 239 GKLA---GMRPIMLELGSKDAGIVLADADLDNAKQIVAGAYDYGQRCTAIKRVLYEE 295
DB 266 KEAASRSNLKRVTLLEGSKNPCIACADLDLVBCHAGVFFNQGQCTTAASRFBVEQ 325
QY 296 VADELAERKISENVAKLSVGDPPDNAVTT-PYIDNSADFIESLVVDARQAK--ELNE 351
DB 326 VYSEFVRSSVYAKRPVGDPEVDTEGQPQIDQKQFKEIILESGKKEGAKJECGSA 385
QY 352 FKRDRLTLPLGFPHVTLDMLAMEEPFGPLPIIRVDAEBAVALNKSDPGIOSSVFT 411
DB 386 MEDKLFIFKTFVFSVTDMRIAKEBELFGVQPLIKFSEIEYIKRAASTVYGLPAAVFT 445
QY 412 RDPQKAPDIANKLEVGTVHINNKTGRGPDNFPPLGKSGAGVQGIIRYSIEAMTANKSV 471
DB 446 KNIDKALALASLEGVTWICNALYA-QAPFGGFMGSGRELGEYALAEYIEVKIVT 504
QY 472 LDM 474
DB 505 IKL 507
```

RESULT 54

AD0143
betaine aldehyde dehydrogenase (EC 1.2.1.8) [imported] - Yersinia pestis (strain C092)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD0143
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Varraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB00001; MUID:21470413; PMID:11586360
A:Accession: AD0143
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <KUR>
A:Cross-references: UNIPROT:Q8ZGV9; GB:AL590842; PIDN:CAQ90007.1; PID:915979229; GSPDB:G
C:Genetics:
A:Gene: betB
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 27.1%; Score 649.5; DB 2; Length 490;

Best Local Similarity 31.8%; Pred. No. 7.6e-34;
Matches 153; Conservative 100; Mismatches 207; Indels 21; Gaps 10;

```
QY 8 YVNGEWSNV--NOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPAAALTYERA 65
DB 10 YINGAVTSTSGDPTFDANPANNGECIADQAAANQDVOKAVAAKQGPVWAAMTAMERS 69
QY 66 QYLAHRAADIIEERDKKEIATVLAKEISKAYNAS-VTEVTRADLRVAEEGIRLSTSD 124
DB 70 RILRAVDILDRNDELAAIETADTGKPLSETRSDIVTGADVLEYVA--GI---IPALE 124
```



```

RESULT 57
P90673
NAD+-dependent betaine aldehyde dehydrogenase (imported) - Escherichia coli (strain O157
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F90673
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gisawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A59629; MUID:21156231; PMID:11258796
A:Accession: F90673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <HAV>
A:Cross-references: UNIPROT:Q8HEC9; GB:BA000007; PIDN:BA033781.1; PID:g13359815; GSPDB:C
C:Genetics:
A:Experimental source: strain O157:H7, substrain R1MD 0509952
A:Gene: EC60358
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match      27.1%; Score 648.5; DB 2; Length 490;
Best Local Similarity 32.2%; Pred. No. 8-34;
Matches 156; Conservative 101; Mismatches 207; Indels 21; Gaps 9;

QY 4 EYQNTVNGEWSKV--QIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALTV 61
DB 6 EQQLTHGGYTSATSGRTETINPANGVTLVQAAGREDVRAVSAQCGKIRAAATA 65
QY 62 YERAOYLKADIIERDKKEIATVLAKESKAYN-ASVTEVVRADLIRYAEGRIST 120
DB 66 MERSRLIRRAVDILERRNDELAKLETLDTGKAYSETSTVDIVTGADVLEYA--GLIPSL 123
QY 121 SADEGKMDASGHLAVIRQPVGIYLAIAPYNPVNLGSKIPALIGVNVMPKPT 180
DB 124 BGSQPLRETS----FVYTRREPLGVAGIGAMNPTOTALMKSAAPALAGNAMI FKBS 179
QY 181 QGSVGLVLAKAFAGLPAAGVFNITGRGSEIGDIYHEHEVNFINTGSPVQRI-- 238
DB 180 VTPPLAKLAELIYSEGLPDGVFNVLPGVGAETGQLTTHPGIAVNSFTGVAAGKAWA 239
QY 239 -GKLAKMRPIMELGKQAGIYLDADLNAKQIVAGAYVDSGQRTAIKRVLEEVA 297
DB 240 NSAASSLKEVTELMELGKSPLIIVFDADLDLADIAMANFSSGQCTGTGTFVPAKCN 299
QY 298 DELAKISENVAKLSVGDPEFNAT--VTPIIDNSADFIESLVVDARQKAKEL---NEK 353
DB 300 AAFEGKILARVERIRAGVDFDPQTNFGLVSPFHRDNLRYLAKGEBGARVLGCGDVLK 359
QY 354 ----RDGRLLTPGLFDHVTLDMLKLAEEBPFGLPIIRYKDAEBAVALANKSDFGLOS 409
DB 360 GYGGLNGAMVAPTVTDCSDDEMTYREEIFGVMSLTLYSEDEVIYRANNDYGLAAGI 419
QY 410 FTRDPKAPADIANKLEVGTVHNNKTRGPDNPFPLGLGSGAGVQGIYSIEAMTNVKS 469
DB 420 VTADNLRARHVRVHQLGAGICWINT-WGESPAEMPVGKSGIGRENGVMTLGSTQVKS 478
QY 470 IYLDLM 474
DB 479 IQVEM 483

RESULT 58
A75608
aldehyde dehydrogenase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: A75608
R:White, O.; Elsen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

```

```

A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75608
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-495 <WHI>
A:Cross-references: UNIPROT:Q9R227; GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AA01223
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0126
A:Map position: 2
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match      27.1%; Score 647.5; DB 2; Length 495;
Best Local Similarity 33.1%; Pred. No. 1e-33;
Matches 159; Conservative 93; Mismatches 199; Indels 29; Gaps 8;

QY 8 YVNGEWSKVQ--IEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALTVYERA 65
DB 26 FIGGEWVTPCGQGLEVNSATEEVMGTVPGETDEARALIAARAADFSWSQTPAEERA 85
QY 66 QYIHKADIIERDKKEIATVLAKESKAYNS-VTEVVRTADLIRYAEGRISTSD 124
DB 86 ALLERTISGLBEROSEIALVAQEVGMPEVLSNVIQGLPA-----VTEWSMAER 135
QY 125 GSKMDASTGHLAVIRQPVGIYLAIAPYNPVNLGSKIPALIGVNVMPKPTQGSV 184
DB 136 IRLPABETLASULLRREPVGVAALTPMNTPLHQIAKAPALAACTVLAKSEVAPL 195
QY 185 SGLVLAKAPADGLPAAGVFNITGRGSEIGDIYHEHEVNFINTGSPVQRIKLAG- 243
DB 196 NAFVLAELIIEAAGAPAGFNLVTGTPVGCVLAAPPEVDVNSFPGSTPAKRSSELAAR 255
QY 244 -MRPIMELGKQDGIYLDADLD---NAKQIVAGAYVDSGQRTAIRVLEEVA 297
DB 256 TVKRALBELGKSPYLITDDDLGSEALVQAVNMGICQCTYNSGQTSALTRMLVPRSL 315
QY 298 DELAKISENVAKLSVGDPEFNAT--VTPIIDNSADFIESLVVDARQKAKEL---ELN 350
DB 316 PEVEAVQAQVVSQVVGSPFDAAITLGLVSDVQREVRGHIHQGVAGARLLINGASAP 375
QY 351 EPRKDRLLTPGLFDHVTLDMLKLAEEBPFGLPIIRYKDAEBAVALANKSDFGLOS 410
DB 376 EGLERGYYVQPTIISDVTPENTIAEEIFGVLVYQPEDEEDVIRINATHYGLAGVW 435
QY 411 TRDPKAPADIANKLEVGTVHNNKTRGPDNPFPLGLGSGAGVQGIYSIEAMTNVKS 470
DB 436 SSDRERAVRARLKTQGVSLNG--GAPNPVAPFGGYKQSGNGRFGRLGLEEFMEVRL 493

RESULT 59
G64878
aldehyde dehydrogenase (NAD) (EC 1.2.1.3) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: G64878; J00397
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
-A.; Rose, D.J.; Mau, B.; Sano, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G64878
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-495 <BLAT>
A:Cross-references: UNIPROT:P23883; GB:AE000228; GB:U00096; NID:G2367117; PIDN:AACT4382.
A:Experimental source: strain K-12, substrain MG1655
R:Heim, R.; Streicher, E.E.
Gene 99, 15-23, 1991
A:Title: Cloning an Escherichia coli gene encoding a protein remarkably similar to mamma
A:Accession: J00397
A:Molecule type: DNA
A:Residues: 1-312, 'R', 314-495 <HEI>

```

A:Cross-references: GB:M88433; NID:g145223; PIDN:AAA23428.1; PID:g145224
A>Note: the authors translated the codon AGG for residue 313 as Ser
C:Genetics:
A:Gene: aldH
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C:Keywords: NAD; nucleotide binding; oxidoreductase; P-loop
F:49-56/Region: nucleotide-binding motif A (P-loop)
F:58-322/Domain: aldehyde dehydrogenase homology <ALDD>
F:267,302/Active site: Glu, Cys #status predicted

Query Match 27.0%; Score 647; DB 1; Length 495;
Best Local Similarity 33.3%; Pred. No. 1,1e-33;
Matches 164; Conservative 95; Mismatches 203; Indels 30; Gaps 11;

```
QY 1 LTKXYQNYVNGEMKSVNQ-LEISPIDSSIGFVPAMTREVVDHAKAGSEALPA-W 56
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 16 LAIENRLFIINGEYTAANEETFEVDPTQAPLAKIARGKSYDIDRALMAAGVFERGDM 75
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 57 AALTYERQYIHKADIERDKEEATYTLAKEISAYNASV-TEVVTATDLIRAAEEG 115
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 76 SLSSPAKRAVANKADIMEAHEALLETDTGKPIRHSIDRDPGARARIRWYAE-- 133
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 116 IRLSTASDE-GGKMDASTGKLAIVIRPOVGIPLAIAPYNPVNLGSKTAPALIGSNV 174
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 134 -----AIDVYGEVATTSHEILAMTVREVGVIATVPNNFPLLCWKLGFPALAGNSV 188
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 175 MEKPEPTQGSVSGLVLAFAEAGLPAGVNTITGRGSEIGDIYVEHEEVNFINFTGSTPV 234
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 169 ILKPEKPSLAIIRLAGLAKENGLDGVLVNTVGTGHEAGQALSRNDIDALAFGSTTT 248
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 235 GQRIGKLAG--MRPIMELGSKDGIYLAAD-DLDNAKQIVAGAYDYGCRCTAIKEV 290
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 249 GKQLKIDGDSMMKVMLEAGSKSANIYFADCPDLCQAASATPAGIFYNQGVCTAGTSL 308
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 221 LVVEEVADELAKISENNVAKLSVGPEDNA-TVTPVIDNSADPIESIYVVDKOKAGKL 349
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 309 LIEESITADEFTALLKQQAQNNQPGHLDPAITTMGTLLIDCAHDSVHSFIREGSKG--- 364
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 350 NEFKRDK-----LITPGIFDHVTLLDMELAMEBPEPILPIIRVDAEAAVALANKSPF 403
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 365 -QLLIDGRNAGLAALIGPIIFVDVDPNASTLSREIFGPVLVTRFISEQALQLANDSQY 423
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 404 GLQSSVYTRDFOKAFDIANKLEVGYTHINNTKGRGPDNPPFLGLKSGAGVGCIRYSIA 463
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 424 GIGAAVMTETDLSPRAHMSRRLKAGSVFVNNYND-GDMTVPFQGYKQSGNGRDKSLHALRK 482
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 464 MTNKAISIVLDMK 475
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 483 FTLEKTIWISLE 494
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

RESULT 60
D69764
benzaldehyde dehydrogenase homolog ycnH - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: D69764
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: D69764
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-462 <KUN>
A:Cross-references: UNIPROT:p94428; GB:Z99106; GB:AL009126; NID:g2632653; PIDN:CAB12199.
A:Experimental source: strain 168
C:Genetics:
A:Gene: ycnH
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.9%; Score 644.5; DB 2; Length 462;
Best Local Similarity 33.5%; Pred. No. 1.5e-33;
Matches 157; Conservative 84; Mismatches 206; Indels 19; Gaps 7;

```
QY 18 NQIELSPIDSSIGFVPAMTREVVDHAKAGREALPMWALTVERAQYIHKADIER 77
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 3 DQLYVNPATGCEIKITIPQASATVEEAIERSHQFRTWSKTSANERTSLKMWELIYE 62
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 78 DKEEATYLAKEISAYNASVTEVVTADLIRAAEEGRILSTADEGKMDASTGKILA 137
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 63 HKEELADLITKENGKPYGEAEVLYGAGYLEMFAEAKRV-----YGRVPAFTGKRI 117
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 138 VIRPOVGIPLAIAPYNPVNLGSKTAPALIGSNVMEKPEPTQGSVSGLVLAFAEAG 197
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 118 VTRQPVGPVVAITPWNFPNMITKAPALAGCTFIIRKAPDTPSLAYELARLAYERG 177
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 198 LPAGVNTITGRGSEIGDIYVEHEEVNFINFTGSTPVQQRIGKLAG--MRPIMELGSKD 255
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 178 IPKDVLYVIGGEIEGNVFTSPKIRITFTGSPVGLKMKNSADVTKVHSMELGHA 237
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 256 AGIYLAADLDNAKQIVAGAYDYGCRCTAIKRLVYVEVADLAKEISENVAKLSVSD 315
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 238 PLIVEDADIDILAVQAAKSKRNAGQTCVCAHRIHESIDEPAKLSSEGVSLKLVGN 297
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 316 PF-DNATVTPVIDNSADPIESIYVVDKOKAGKL--NEFKRDK-----RLITPGIFDHVT 368
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 298 GLIEEVNVPILNKRGFEKITYSIDAVKAKVAGCTYDRNDKCGCFVNPVLTVDV 357
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 369 LDMKLAMEBPEPILPIIRVDAEAAVALANKSPDFLOSVTRDFOKAFDIANKLEVGT 428
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 358 TSMNIMHEETFGVPVPIYTFSDIDBALQLANDTPYGLAAVFTENYRRGIRYISEMLEYGI 417
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 422 VAHNKTKGRPD--NPPFLGLKSGAGVGCIRYSIEAMTNVSIYVDM 474
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 418 IGMND---GGPVAQAPPGMKESGIGREGESEGIPELYETKYLISGL 462
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

RESULT 61
C69814
benzaldehyde dehydrogenase homolog yfMT - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: C69814
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: C69814
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-465 <KUN>

A:Cross-references: UNIPROT:O06478; GB:Z99107; GB:Z99108; GB:AL009126; NID:G2633055; PID
A:Experimental source: strain 168
C:Genetics:
A:Gene: yfnt
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.9%; Score 643.5; DB 2; Length 485;
Best Local Similarity 33.5%; Pred. No. 1.8e-33;
Matches 158; Conservative 86; Mismatches 216; Indels 11; Gaps 5;

QY 6 QNYVNGEMK--SSVNOIELSPIDSSLGFPVAMTBEEDVHAMKAGREALPMAALTYYE 63
DB 9 KQIFGKQEGSSPVLENKNRPYKQFTTFRKATADVDYRAALAKKKMDVNPPE 68
QY 64 RAQYLHKADILIERKEEIVATLAKEISKAVNASVTEVVTADLIRYAESEGRISTSD 123
DB 69 KRTLEKAVTYIEENBEALITYIMELGSTRKKAPEIGLVKNITKEAATPIRM----- 123
QY 124 EGGKNDASTGHLAVIRROPVIGVIALADYNYPNVLSGSKIAPALIGNVNFKPPTQGS 183
DB 124 ESKILPSTIDGKENLRYFPAVGVGVISPFNPFSLKMSVAPALGAGNGVVLKPHETP 183
QY 184 V-SGLVAKAFAGAPAGVNTITGRSGEIGDIYVEHEVNFINTGSTPYGQRIGKA 242
DB 184 ICGTLLAKIFENAGIPAGLNVVTDIAEIGDSFVEHPRIISTSTKVGSIYIGLA 243
QY 243 --GMRPIMLELGKDGAGIVLADADLDNAKQIVAGAYDYSGORCTAIKRVLVVEVDEL 300
DB 244 MKHFKKPILLEGNSAFIYLEDADIEYAVNAVFSRFTHQGLCMSANRVLVHSSLYDKF 303
QY 301 AEKISENVAKLSVGPDPFNATVT-PVIDNSADPIESLVVDARQKAKELNEFKRDRLL 359
DB 304 LELYQAKVESLGVGPMPDITIIIGPLINSRQTDGMKTYEQAIIEGAVPVKLGGFNGTIV 363
QY 360 TPGLDVHTLDMKLAMEBEPGPIPIIRKDAEENAIKNSDPGQSSVFRDQKAD 419
DB 364 EBTILKDVAFWISAKIELGSPVSMKDSDEADVANIETPFGISGVHNSIRGEVA 423
QY 420 IANKLEVTGVIHNNKTRGPDNPFGLKSGAGVQGIYRSIEAMTNVXSI 470
DB 424 FAKRIETGMIHVNDTINDERVAVGEGKQSLGRANGMSLEEFITLKMI 474

RESULT 62
C83675
glycine betaine aldehyde dehydrogenase gbsa [imported] - Bacillus halodurans (strain C-1
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: C83675
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: AB3650; MUID:20512582; PMID:11058132
A:Accession: C83675
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <STO>
A:Cross-references: UNIPROT:O9KGA9; GB:AP001507; GB:BA000004; NID:G10172612; PIDN:BA8039
A:Experimental source: strain C-125
C:Genetics:
A:Gene: gbsa
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.9%; Score 643.5; DB 2; Length 490;
Best Local Similarity 32.9%; Pred. No. 1.8e-33;
Matches 161; Conservative 89; Mismatches 200; Indels 39; Gaps 8;

QY 5 YONVNGEMKSSVNOIELSPIDSSLGFPVAMTBEEDVHAMKAGREALP--AMALTY 62
DB 6 YMVVNGERVDSDPEFTENPAKGEILATVAKASRHNHAKVQAAMHAPDHQKWKYRPG 65
QY 63 EAAQVLAADILIEEDKEIATVLAKISKAVNASVTEVVTADLIRYAAEEGIRLSTSA 122

DB 66 KRATVNOIAIMBERFOELVEIVLNSKAISAQAQINQAIIDFEFFYA----- 115
QY 123 DEGGKNDASTGHK-----LAVIRROPVIGVIALADYNYPNVLSGSKIAPALIGSN 172
DB 116 -----GAVHGRGVNNVNPGLFANTYQKEVGVCAQIIPNYPILMMAMKAVAPALAGC 169
QY 173 VVMFKPPTQGSVGLVAKAFAGAPAGVNTITGRSGEIGDIYVEHEVNFINTGST 232
DB 170 SVVVPASLITPITAILNEICHEAGVPEGVNNTVAGASVVDYLVGHEVDNVKVAFTST 229
QY 223 PVGQRIGKLAG--MRPILELGKDGAGIVLADADLDNAKQIVAGAYDYSGORCTAIKRV 290
DB 220 PIGKDIMKASHTLKRVTLEIGGKSPNIVFADADMEAAVAGSLRGITYNPTQSCARSRL 289
QY 291 LVVEVDELAKESIVAKLSVGPDPFNAT-VTVIDNSADPIESLVVDARQKGA--- 346
DB 290 FVHEIYIEFIEFTTRAKKIQLEDPPFKTHVGSIIIRDQLEITDSYVRSAEEGAIYA 349
QY 347 ---KELN-EFKRDGELLTPGLFDVHTLDMKLAMEBEPGPIPIIRKDAEENAIKNSD 402
DB 350 LGKERRVEGEPNGHWYEPVITVNTVPDKAVQSEIFGPVVVVEFPQDEREVIRKANDTK 409
QY 403 FGLQSVTRDPQKAFDLANKLEVCTHAIHNNKTRGPDNPFGLKSGAGVQGIYRSIE 462
DB 410 FGLSALWTKDQGRATRVAHQLEAGIWMVNCPSAPPEP-DPGYKESGFRELCVETLD 468
QY 463 AMTNVKSIV 471
DB 469 LYMERKSL 477

RESULT 63
E90863
aldenhyde dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain R1MD 050
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: E90863
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gssawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A9629; MUID:21156231; PMID:11258796
A:Accession: E90863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-495 <HAY>
A:Cross-references: UNIPROT:O8XTG6; GB:BA000007; PIDN:BA835300.1; PID:G13361342; GSPDB:G
A:Experimental source: strain O157:H7, substrain R1MD 050952
C:Genetics:
A:Gene: BC81877
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.9%; Score 643; DB 2; Length 495;
Best Local Similarity 33.1%; Pred. No. 2e-33;
Matches 163; Conservative 96; Mismatches 203; Indels 30; Gaps 11;

QY 1 LTKEXONVNGEMKSSVNO--IELISPIDSSLGFPVAMTBEEDVHAMKAGREALP--W 56
DB 16 LAIENRPLFNGEYTLAAENETPEYDPTQAPLANIANGKSVDDIDRAVSARGVBERDW 75
QY 57 AALTYERAOYLHKADILIERKEEIVATLAKEISKAVNASV-TEVVTADLIRYAAEG 115
DB 76 SLSPSPAKKAVLNLKADILIEANAEIALLETLDCKPIRHSILRBDIPQAAAIKRYAE-- 133
QY 116 IRLSTSADE-GGKNDASTGHLAVIRROPVIGVIALADYNYPNVLSGSKIAPALIGNAV 174
DB 134 ----AIDKVGGEVATTSHEILAMVREPVGIAIYVWNPPLLTCTCKGLBPALAAAGSV 188
QY 175 MKPPTQGSVGLVAKAFAGAPAGVNTITGRSGEIGDIYVEHEVNFINTGSTPV 234
DB 189 VAKPEKSPLSHIRLAGAKERGLPDGLNVVTVGHEAGQSLSHNDIDIAFTGSTRT 248
QY 235 GQRIGKLAG--MRPIMLELGKDGAGIVLADADLDNAKQIVAGAYDYSGORCTAIKRV 290

RESULT 71

aldehyde dehydrogenase homolog - Sphingomonas aromaticivorans plasmid pNL1
T31293
C:Species: Sphingomonas aromaticivorans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T31293
R:Romano, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G
submitted to the EMBL Data Library, July 1998
A:Description: Complete sequence of a 184 kb catalytic plasmid from Sphingomonas aromati
A:Reference number: Z20992
A:Accession: T31293
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-479 <ROW>
A:Cross-references: UNIPROT:O86001; EMBL:AF079317; NID:93378261; PID:93378434; PIDN:AAD
C:Genetics:
A:Genome: plasmid pNL1
A>Note: nahf
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.6%; Score 637; DB 2; Length 479;

Best Local Similarity 33.8%; Pred. No. 4, 6e-33;

Matches 153; Conservative 75; Mismatches 217; Indels 8; Gaps 3;

QY 23 LSPIDSSIGFVPAMTREFVDHAKGREALPAMALTYVERAQLYHKAADIIRDKERI 82

DB 24 LNPVDSVYSRVPAGAEVNAVENAHQYLOHRLPAAVBEGLAKAEIEMERTAF 83

QY 83 ATFLAKEISKAYNAVSTEVVTRADLIRVAEGRISTDSBGKMDASTGKLAIVIRQ 142

DB 84 ADVLDELISPIAKGFETRFVAFSLRAAIGVPRRI-----RGETIPSDTPGRFSMLIQ 138

QY 143 PGIYLAIPVYVNVLSGSKTAPALIGNVVMFKPTQGSGLVLAAPAEAGPAPV 202

DB 139 PGGVAGITPVPVPLIKGKQSMALATNAPVLAPEAPMADLALMKRAGVPPDL 198

QY 203 FNTITRGSEIDYIYEHEVNFINTGSPVQORIGKLA--GMRPIMLELGGKDGIVL 260

DB 199 FNVVNGNGEIDVLTGHRKVASITTTGSSRYGKHAETIAARKTKYTELEGGKSLVVC 258

QY 261 ADADLDNAKQIVAGAYDYSQRCIAIKVLVEEVADELAEKISENVAKLSGVPFDNA 320

DB 259 ADADLDKAVNALFSIFMVGQVCMGASRIYVERSIFDQFTAFAPAAATGRANSGLDRDPT 318

QY 321 TYT-FVIDNSADFTIESLVVDAROKAKALNEFKRDGLITLPHVTLDMKLAMEEPF 379

DB 319 TMLGPISRRQDRVRRHIDDAKSKAAVLAGGEMSGNSCAATILSGVTAEMTVFEETTF 378

QY 380 GPILPIIRVKDAEBAVALANKSDFGLOSVPFRDFQKAPDIANKLEVGTVHINNKTRGRP 439

DB 379 GVTSLFFPDTIEBELNNTNTEYGLSASIFTRDLKALAFAPORABGVHINAPTLHDE 438

QY 440 DNFPLGLKSGAGVQGIYSIEMNTVKSIVL 472

DB 439 PHVPGGTAKASGFGREGTEADLEIMTEMKWVTI 471

RESULT 72

F98208

sucinate-semialdehyde dehydrogenase PA0265 [imported] - Agrobacterium tumefaciens (stra

C:Species: Agrobacterium tumefaciens

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C:Accession: F98208

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wolim, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2333-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: F98208

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-484 <KUR>

A:Cross-references: UNIPROT:O8U852; GB:AE007870; PIDN:AAK89192.1; PID:g15159010; GSPDB:GR

C:Genetics:

A:Gene: AGR L 1228

A:Map position: linear chromosome

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.6%; Score 636; DB 2; Length 484;

Best Local Similarity 33.2%; Pred. No. 5, 4e-33;

Matches 157; Conservative 82; Mismatches 218; Indels 16; Gaps 7;

QY 9 VNGEKSSVNO--IEILSPIDSSIGFVPAMTREFVDHAKGREALPAMALTYVERAQ 66

DB 15 VAGRWIGCDDEBTATIRNPATGKIVQVPEIGATEQTQAITPAVAQAKMARRTGERAA 74

QY 67 YLHKADIIIEEDKEIALAVLAKESKANNAVSTEVVTRADLIRVAEGRISTDSBG 126

DB 75 ILKANHRLVMEKNDLGMILTLLEGKPLAEKAGIAYGASFLEWFAEARRRINGETVPQH 134

QY 127 KMDASTGKHLAVIRROPVIGVLAIPVYVNVLSGSKTAPALIGNVVMFKPTQGSVSG 186

DB 135 QAD-----KRLIVLRQPGVAAIIPVNPFPNAMIIRKGPALAGCAVVLKAPDTPESA 189

QY 187 LVLAAPAEAGLPACVNTITRGSEIDYIYEHEVNFINTGSPVQORIGKLAG--M 244

DB 190 IALAILARAGIPLELFSIVTGPAAIEIGVLTASPDILTLFTGSTRGBHLYRCCAPTI 249

QY 245 PRIMELEGKQAGIYADADLDNAKQIVAGAYDYSQRCIAIKVLVEEVADELAEKI 304

DB 250 KGLGLELGNNAFIFENDADLDAAVEGALIAKFRNNGTCCANFLYQSGVYDFAEKF 309

QY 305 SENVAKLSVGPFDNATVT-FVIDNSADFTIESLVVDAROKAKEL--NEFKRDGLLT 360

DB 310 ARAVSGLVKNGLEBSITLGPIDGNVAKYBAHIALDSGAQIAAGSRYALGPNFE 369

QY 361 PGLPHVTLDMKLAMEEPPGPIPIIRVKDAEBAVALANKSDFGLOSVPFRDFQKADI 420

DB 370 PTLIRNVADMQVAEEFTGPIALPLFREDESDVIAQANNDFGIASYFYARDLSRVRV 429

QY 421 ANKLEVGTVHINNKTR--GPDNFPFLGLKSGAGVQGIYSIEMNTVKSIVL 472

DB 430 AEALSYGVGVN--TGAISTABAPFGVGMKSLGREGSRHGLEEYTELKVCYCI 480

RESULT 73

AC3078

aldehyde dehydrogenase atck [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C:Accession: AC3078

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClellan,

A.; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.

ster, B.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AC3078

A:Status: preliminary

A:Molecule type: DNA

A:Cross-references: UNIPROT:O8U852; GB:AE008689; PIDN:AAL45041.1; PID:g17742704; GSPDB:GR

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: atck

A:Map position: linear chromosome

C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.6%; Score 636; DB 2; Length 484;

Best Local Similarity 33.2%; Pred. No. 5, 4e-33;

Matches 157; Conservative 82; Mismatches 218; Indels 16; Gaps 7;

```

QY 9 VNGEMKSVNO--IEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVERA 66
DB 15 VAGRMIGCDRETATIRNPAATGKIYQVVELGATEFOALITAAVIAOKMARTAGERAA 74
QY 67 YLHKADIIERDEKEIATVLAKEISKAVNASVTEVVRTADLIRYAAEGIRLSTADEG 126
DB 75 ILKAMRLVEMNRDDELGMILLTEQCKPLAEAGELIYVGSFIEMFEEARIRINGEVRPH 134
QY 127 KMDASTGHLAVITROPVGIYLAIAFYNTPNLISGSKIAIPALIGNVVMFKPTQGSVG 186
DB 135 QAD-----KRIIVLRQPVGVAAITPMNPPNMITKRVGPALAAAGCAVVLKAPQPFPA 189
QY 187 LVLAFAFAAGLPAQVFNITGRSGEIRGYIYHEHEVNFINTGSPVQORIGKLAG--M 244
DB 190 IMLAIIAERAGLPLEFISITVGPMAEIGVLTASPIRLITFTGSTRTEBHLRQCAPTI 249
QY 245 RPIMLELGGKAGIYVADADLDNAAKQIVAGAVDYGQRCTAIKRYLVVEVADELAETK 304
DB 250 KTLGELGGMNAPFIVENDADLDAVVGALIAKFRNNGQCTVCANRLYVQSGYDAFAEKF 309
QY 305 SENYAKLSVGPDPDNATVT--PVIDNSADPIESLVVDARQKAKEL--NEFKRGRLLT 360
DB 310 ARAVSGILKVGNGLEBESIIIGPLIDGNAVAKVEAHIDALSKGQIAAGSRYALGNGFPE 369
QY 361 PGLFPHVTLDMKLAEEPPGPIIPIRYKDAEAAVALANKSDPGLQSSVPTDFQKAPFI 420
DB 370 PTLIRNVITAMOVARETGPPLALPLRFREDEEDVIAQANNTPGLASIFYARDLSRVFRV 429
QY 421 ANKLEGVTHINNKTR--GPDNFPPLGLKSGAGVQIGIRYSIEAMTNVKSIVL 472
DB 430 ABALEVGWGVN--TGAISTAEKPFQGVMSGLGRGSHGIEFYELKYYVCI 480

RESULT 74
S74224
aldehyde dehydrogenase (NAD) (EC 1.2.1.3) 2 - mouse
N.Alternate names: retinaldehyde-specific dehydrogenase
C.Species: Mus musculus (house mouse)
C.Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C.Accession: S74224
R.Zhao, D.; McCaffery, P.; Ivans, K.J.; Neve, R.L.; Hogan, P.; Chin, W.W.; Draeger, U.C.
Eur. J. Biochem. 240, 15-22, 1996
A.Title: Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinaldehyde dehydrogenase
A.Reference number: S74224; MUID:96390857; PMID:8797830
A.Accession: S74224
A.Molecule type: mRNA
A.Residues: 1-499 <ZHA>
A.Cross-references: UNIPROT:Q62148; EMBL:X89273; NID:g1430868; PIDN:CA67666.1; PID:g1430868
A.Experimental source: strain C3H/HeJ; Cell type embryonal carcinoma; cell line P19 teratocarcinoma
C.Gene: RALDH-2
C.Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
C.Keywords: NAD; oxidoreductase
F:57-321/Domain: aldehyde dehydrogenase homology <ALDH>
F:193-273/Domain: NAD binding #status predicted <NAD>
F:267-301/Active site: Glu, Cys #status predicted
F:454/Binding site: NAD (Cys) #status predicted

Query Match 26.5%; Score 635; DB 2; Length 499;
Best Local Similarity 33.5%; Pred. No. 6, 5e-33;
Matches 161; Conservative 84; Mismatches 210; Indels 26; Gaps 9;

```

```

QY 179 PTQGSVGLVLAFAEAGLPAGVFNITGRSGEIRGYIYHEHEVNFINTGSPVQORI 238
DB 193 AEOTPLSALYMGALIKAGFPFGVNNILPGYPTAGAAASHIGIDKIAFTSTVEGKLI 252
QY 239 GKLAG--MRPIMLELGGKAGIYVADADLDNAAKQIVAGAVDYGQRCTAIKRYLVVE 295
DB 253 QFAAGRSNLKRYTLELGGKSPNITFADADLDVAEDAHQGVFPNQGCTAASRIIVESS 312
QY 296 VADELAEKISENVAKLSVGPDPDNATVT--PVIDNSADPIESLVVDARQKAKELNEFK- 353
DB 313 IYEEFVKRSVERAKRIVGSPEDPTTEQGPQIDKQYNNVLELIGSVAEKAKECGKGK 372
QY 354 --RDGRLLTPGLFDVHTLDMKLAEEPPGPIIPIRYKDAEAAVALANKSDPGLQSSVPT 411
DB 373 LGRKGFPEFPVFSNVTDMDRIAKEBIFGPVQEIIRFKTMDVIERANNSDFGLVAAPFT 432
QY 412 RDPKAPDIANKLEGVTHINNKTRGPDNFPPLGLKSGAGVQIGIRYSIEAMTNVKSIV 471
DB 433 NDINKALVSSAMQGTWINCYNALMAQS--PFGGFKMSGNGRNGEFGIREYSEKIVT 491
QY 472 L 472
DB 492 V 492

RESULT 75
H98159
hypothetical protein AGR_L1010GM [imported] - Agrobacterium tumefaciens (strain C58, Ce
C.Species: Agrobacterium tumefaciens
C.Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C.Accession: H98159
R.Goodner, B.; Hinkle, G.; Galtung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A.Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A.Reference number: A97359; MUID:21608551; PMID:11743194
A.Accession: H98159
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-483 <KUR>
A.Cross-references: UNIPROT:Q8U711; GB:AE007870; PIDN:AKK88802.1; PID:g1515855; GSPDB:
C.Gene: AGR_L1010GM
A.gene: AGR_L1010GM
A.Map position: linear chromosome
C.Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 26.4%; Score 632.5; DB 2; Length 483;
Best Local Similarity 34.0%; Pred. No. 9e-33;
Matches 161; Conservative 89; Mismatches 210; Indels 13; Gaps 5;

```

Db 302 VEAFAKKAASLTAGDPRECKTEPLGSLVSAEAASRIVLVDVAVSKARRVAGGGDTML 361
Qy 360 TPGLFHDYTLDMKCLAMEEPFGPILPIIRYKDAEAAVATANKSDPGLQSSVFTTRDFOKAFD 419
Db 362 DAIAVDGVTTPARLYSEESFGPVSIIRAGSTIDEAVSIAESEFGLSAAVFGRDQARALS 421
Qy 420 IANKLEVSTVHINNKTGRGPDNFPFLGLKSGAGVQGI RYSIEAMTNVKSIVL 472
Db 422 VAARIESGICHVNGPTVHDEAQLPFGGVKASGVGRFGGTAGIAEFTELRWTL 474

Search completed: August 28, 2005, 10:28:32
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2005, 10:05:09 ; Search time 176 Seconds

(without alignments)
1382.032 Million cell updates/sec

Title: US-09-868-195-12

Perfect score: 2393
Sequence: 1 LTRKQNTYNGEWSVNOI.....GIRYSLEANTNYSIVLDMK 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 80 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2391	99.9	475	2	Q8RP84 streptococc
2	2382	99.5	475	2	Q8E0B1 streptococc
3	2379	99.4	475	2	Q8E5Y8 streptococc
4	1812	75.7	474	2	Q97Q70 streptococc
5	1802	75.3	474	2	Q8DPS7 streptococc
6	1700	71.0	475	1	GAPN STRMU
7	1645	68.7	475	2	Q8P0F5 streptococc
8	1639	68.5	475	2	Q99267 streptococc
9	1618	67.6	467	2	Q8K707 streptococc
10	1383	57.8	482	2	Q97D25 clostridium
11	1345	56.2	479	2	Q73CX4 clostridium
12	1344	56.2	479	2	Q6HMX0 bacillus th
13	1341	56.0	479	2	Q63FF8 bacillus th
14	1338	55.9	479	2	Q81UL6 bacillus an
15	1335	55.8	477	2	Q72Z26 bacillus ce
16	1333	55.7	479	2	Q81HE6 bacillus ce
17	1333	55.3	482	2	Q8XHP4 clostridium
18	1244	52.0	481	2	Q9KXN0 bacillus ha
19	1171.5	49.0	496	2	Q9SNX8 apium grave
20	1168.5	48.8	496	2	Q9ZUG8 arabidopsis
21	1164.5	48.7	498	1	GAPN MAIZE
22	1161.5	48.5	496	1	GAPN_PEA
23	1158.5	48.4	499	2	Q6ZNG0
24	1155.5	48.3	496	1	GAPN NICPL
25	1145.5	47.9	507	2	Q8S4Y9
26	1138.5	47.6	508	2	Q84PC4
27	1120.5	46.8	496	2	Q6RCS1
28	1079	45.1	471	2	Q6MT99
29	1077.5	45.0	472	2	Q6FLK7 mesoplasma
30	1047.5	43.8	478	2	Q8EVT9 mycoplasma
31	1027.5	42.9	496	2	Q8LKE1 triticum ae

32	1026.5	42.9	496	2	Q8L5J9	Q8L5J9 triticum ae
33	994	39.4	525	2	Q7NBX5	Q7NBX5 mycoplasma
34	868.5	36.3	475	2	Q9PQC9	Q9PQC9 ureaplasma
35	810.5	33.9	475	2	Q7A5E6	Q7A5E6 geobacter s
36	790	33.0	478	2	Q75TD2	Q75TD2 geobacillus
37	786	32.8	478	2	Q8ZVJ1	Q8ZVJ1 pyrobaculum
38	763	31.9	488	1	DHAL_BACST	P42236 bacillus su
39	759.5	31.7	463	1	YELL_METUA	Q58806 methanococ
40	752	31.4	470	2	Q97XAS	Q97XAS sulfolobus
41	751	31.4	468	2	Q976X5	Q976X5 sulfolobus
42	750.5	31.4	486	2	Q8EMK2	Q8EMK2 oceanobacil
43	750	31.3	470	2	Q97X59	Q97X59 sulfolobus
44	742	31.0	497	2	Q8CV96	Q8CV96 oceanobacil
45	741.5	31.0	455	2	Q27059	Q27059 methanobact
46	741.5	31.0	484	2	Q8NMB0	Q8NMB0 corynebacte
47	741.5	31.0	489	2	Q9HUR4	Q9HUR4 pseudomonas
48	741.5	31.0	496	2	Q6M2H6	Q6M2H6 corynebacte
49	739.5	30.9	488	2	Q65NX0	Q65NX0 bacillus th
50	739	30.9	506	2	Q8U2S5	Q8U2S5 pyrococcus
51	737	30.8	476	2	Q8E573	Q8E573 aquifex aeo
52	736.5	30.8	491	2	Q59702	Q59702 pseudomonas
53	736	30.8	474	2	Q81QX6	Q81QX6 bacillus an
54	736	30.8	474	2	Q6HJ76	Q6HJ76 bacillus th
55	735	30.7	474	2	Q738S1	Q738S1 bacillus ce
56	735	30.7	482	2	Q9KC36	Q9KC36 bacillus ha
57	734	30.7	471	2	Q8TWC7	Q8TWC7 methanopyru
58	732	30.6	474	2	Q63B07	Q63B07 bacillus ce
59	726	30.3	482	1	GABD_ECOLI	P25526 escherichia
60	726	30.3	488	2	Q721Q0	Q721Q0 listeria mo
61	725	30.3	499	2	Q8RFU1	Q8RFU1 corynebacte
62	722.5	30.2	483	2	Q73EK2	Q73EK2 bacillus ce
63	721.5	30.2	483	2	Q81ZE2	Q81ZE2 bacillus an
64	721.5	30.2	483	2	Q6HP75	Q6HP75 bacillus th
65	721.5	30.2	488	1	DHAL_BACST	P42232 bacillus in
66	720	30.1	488	2	Q92DA3	Q92DA3 listeria st
67	718.5	30.0	483	2	Q63G03	Q63G03 bacillus ce
68	718	30.0	488	2	Q8Y8I9	Q8Y8I9 listeria ce
69	717	30.0	482	2	Q8X950	Q8X950 escherichia
70	716.5	29.9	493	2	Q9HK01	Q9HK01 thermoplasma
71	716	29.9	474	2	Q81DV8	Q81DV8 bacillus ce
72	713.5	29.8	480	2	Q9HQZ2	Q9HQZ2 halobacteri
73	713	29.7	501	2	Q57693	Q57693 thermoprote
74	711.5	29.7	472	2	Q8EMH9	Q8EMH9 oceanobacil
75	711.5	29.7	483	2	Q81LM7	Q81LM7 bacillus ce
76	711	29.7	475	2	Q8BLI8	Q8BLI8 oceanobacil
77	710.5	29.7	489	2	Q9HMT6	Q9HMT6 halobacteri
78	710.5	29.7	498	2	Q8XOP9	Q8XOP9 anabaena sp
79	710	29.7	486	2	Q7VSP3	Q7VSP3 bordetella
80	710	29.7	486	2	Q7W373	Q7W373 bordetella

ALIGNMENTS

RESULT 1
ID Q8RP84 PRELIMINARY; PRT; 475 AA.

AC Q8RP84; 01-JUN-2002 (TREMBLER, 21, Created)
DT 01-JUN-2002 (TREMBLER, 21, Last sequence update)
DT 01-MAR-2004 (TREMBLER, 26, Last annotation update)
DE Non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1311;
RX MEDLINE=21643113; PubMed=11854208;
DOI=10.1128/IAI.70.3.1254-1259.2002;
RA Hughes M.J., Moore J.C., Lane J.D., Wilson R., Pribul P.K.,
RA Younes Z.N., Dobson R.J., Everest P., Reason A.J., Redfern J.M.,
RA Greer F.M., Paxton T., Panico M., Morris H.R., Feldman R.G.,

RA Santangelo J.D.;
 RT "Identification of major outer surface proteins of *Streptococcus*
 agalactiae.";
 RL Infect. Immun. 70:1254-1259(2002).
 DR EMBL: AF439646; AAL85685.1; -.
 DR HSSP: O59931; 1E0H.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehyd.rog.
 DR Pfam: PF00171; Aldehyd. 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
 SQ SEQUENCE 475 AA; 51198 MW; C525151C88B6AC01 CRC64;

Query Match 99.9%; Score 2391; DB 2; Length 475;
 Best Local Similarity 99.8%; Pred. No. 5.5e-135;
 Matches 474; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTKETONTYNGEWSKSSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPMAALT 60
 :|||||
 DB 1 MTKETONTYNGEWSKSSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPMAALT 60
 QY 61 VYERAOYLHKADIIERDKETATYLAKEISKAYNASVTEVTRADLIRYAABEGIRLT 120
 :|||||
 DB 61 VYERAOYLHKADIIERDKETATYLAKEISKAYNASVTEVTRADLIRYAABEGIRLT 120
 QY 121 SADEGKMDASTGHLAVIRROPVGIIVLAIPYNPVNLSSGKIAPALIGNVVMFKPPT 180
 :|||||
 DB 121 SADEGKMDASTGHLAVIRROPVGIIVLAIPYNPVNLSSGKIAPALIGNVVMFKPPT 180
 QY 121 SADEGKMDASTGHLAVIRROPVGIIVLAIPYNPVNLSSGKIAPALIGNVVMFKPPT 180
 :|||||
 DB 121 SADEGKMDASTGHLAVIRROPVGIIVLAIPYNPVNLSSGKIAPALIGNVVMFKPPT 180
 QY 181 QGSVSGVLAKAPAEAGLPAGVFNITITGRGSEIGDIYVEHEEVNFINTGSTPVGQRIK 240
 :|||||
 DB 181 QGSVSGVLAKAPAEAGLPAGVFNITITGRGSEIGDIYVEHEEVNFINTGSTPVGQRIK 240
 QY 241 LAGMPPIMLGKDGAGIYLAADADLNAAKOIVAGAYDSGORCTAKRVIVVEEVADEL 300
 :|||||
 DB 241 LAGMPPIMLGKDGAGIYLAADADLNAAKOIVAGAYDSGORCTAKRVIVVEEVADEL 300
 QY 301 AEKISENVAKLSVGPDPFNATVTPVIDNSADFIESTLVVDAROKAKELNEFKDGRILT 360
 :|||||
 DB 301 AEKISENVAKLSVGPDPFNATVTPVIDNSADFIESTLVVDAROKAKELNEFKDGRILT 360
 QY 361 PGLFDHVTLDKMLAEPEFGPILPIIRYKDAEBAVAIANKSDFGIQSSVFTTRDFOKAFDI 420
 :|||||
 DB 361 PGLFDHVTLDKMLAEPEFGPILPIIRYKDAEBAVAIANKSDFGIQSSVFTTRDFOKAFDI 420
 QY 421 ANKLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475
 :|||||
 DB 421 ANKLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475

RESULT 2

Q8E0B1 PRELIMINARY; PRT; 475 AA.
 AC Q8E0B1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 GN Glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent.
 OS Name=gagn; Ordered locus names=SA0823;
 OC Streptococcus agalactiae (serotype V).
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCB1_Taxid=216466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2603 V/R / Serotype V;
 RX MEDLINE=22222388; PubMed=12200547; DOI=10.1073/pnas.182380799;
 RA Tetselin H., Mastignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,
 RA Wessels M.R., Paulsen I.T., Nelson K.B., Margalit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,

RA Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
 RA Rinaldo D., Rappapoli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 emerging human pathogen, serotype V *Streptococcus agalactiae*,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 DR EMBL: AB014229; AAM99710.1; -.
 DR HSSP: O59931; 1E0H.
 DR TIGR: SAC0823; -.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehyd.rog.
 DR Pfam: PF00171; Aldehyd. 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
 KW Complete proteome.
 SQ SEQUENCE 475 AA; 51218 MW; F91C086696BAE668 CRC64;

Query Match 99.5%; Score 2382; DB 2; Length 475;
 Best Local Similarity 99.2%; Pred. No. 1.9e-134;
 Matches 471; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTKETONTYNGEWSKSSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPMAALT 60
 :|||||
 DB 1 MTKETONTYNGEWSKSSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPMAALT 60
 QY 61 VYERAOYLHKADIIERDKETATYLAKEISKAYNASVTEVTRADLIRYAABEGIRLT 120
 :|||||
 DB 61 VYERAOYLHKADIIERDKETATYLAKEISKAYNASVTEVTRADLIRYAABEGIRLT 120
 QY 121 SADEGKMDASTGHLAVIRROPVGIIVLAIPYNPVNLSSGKIAPALIGNVVMFKPPT 180
 :|||||
 DB 121 SADEGKMDASTGHLAVIRROPVGIIVLAIPYNPVNLSSGKIAPALIGNVVMFKPPT 180
 QY 121 SADEGKMDASTGHLAVIRROPVGIIVLAIPYNPVNLSSGKIAPALIGNVVMFKPPT 180
 :|||||
 DB 121 SADEGKMDASTGHLAVIRROPVGIIVLAIPYNPVNLSSGKIAPALIGNVVMFKPPT 180
 QY 181 QGSVSGVLAKAPAEAGLPAGVFNITITGRGSEIGDIYVEHEEVNFINTGSTPVGQRIK 240
 :|||||
 DB 181 QGSVSGVLAKAPAEAGLPAGVFNITITGRGSEIGDIYVEHEEVNFINTGSTPVGQRIK 240
 QY 241 LAGMPPIMLGKDGAGIYLAADADLNAAKOIVAGAYDSGORCTAKRVIVVEEVADEL 300
 :|||||
 DB 241 LAGMPPIMLGKDGAGIYLAADADLNAAKOIVAGAYDSGORCTAKRVIVVEEVADEL 300
 QY 301 AEKISENVAKLSVGPDPFNATVTPVIDNSADFIESTLVVDAROKAKELNEFKDGRILT 360
 :|||||
 DB 301 AEKISENVAKLSVGPDPFNATVTPVIDNSADFIESTLVVDAROKAKELNEFKDGRILT 360
 QY 361 PGLFDHVTLDKMLAEPEFGPILPIIRYKDAEBAVAIANKSDFGIQSSVFTTRDFOKAFDI 420
 :|||||
 DB 361 PGLFDHVTLDKMLAEPEFGPILPIIRYKDAEBAVAIANKSDFGIQSSVFTTRDFOKAFDI 420
 QY 421 ANKLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475
 :|||||
 DB 421 ANKLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDMK 475

RESULT 3

Q8E5Y8 PRELIMINARY; PRT; 475 AA.
 AC Q8E5Y8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 GN Hypothetical protein gbs0841.
 OS Ordered locus names=gbs0841;
 OC Streptococcus agalactiae (serotype III).
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCB1_Taxid=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEW316 / Serotype III;
 RX MEDLINE=22242508; PubMed=12354221;

RA Glaeser P., Rusniok C., Buchrieser C., Chevallier F., Frangoul L.,
RA Maadik T., Zouine M., Couve E., Lailou L., Poyat C., Trieu-Cuot P.,
RA Kunst F.;
RT "genome sequence of *Streptococcus agalactiae*, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766847; CAD46485.1; -.
DR HSSP; Q59931; IEUH.
DR Sagalier; sbe0841; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR Complete proteome.
DR KW complete proteome.
SQ SEQUENCE 475 AA; 51217 MW; DE3AE79B81A66B6 CRC64;

Query Match 99.4%; Score 2379; DB 2; Length 475;
Best Local Similarity 99.2%; Pred. No. 2.9e-134; Indels 0; Gaps 0;
Matches 471; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTKYQNYVNGEKSSVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALYT 60
DB 1 MTKYQNYVNGEKSSVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALYT 60
QY 61 VYRAOYLHKADIIERDKKEIATVLAKEISKAYNASVTEVATADLIRYAAEEGIRLST 120
DB 61 VYRAOYLHKADIIERDKKEIATVLAKEISKAYNASVTEVATADLIRYAAEEGIRLST 120
QY 121 SABEGGMASSTGHKLAVIRORPGIVLAIPAPYNPVNLGSKIAPALIGGNVMPKPPT 180
DB 121 SABEGGMASSTGHKLAVIRORPGIVLAIPAPYNPVNLGSKIAPALIGGNVMPKPPT 180
QY 181 QGSVSGVLVAKAEAGLPAQVFNITGRGSEIGDYVEHEEVNFINTGSTPVGORIGK 240
DB 181 QGSVSGVLVAKAEAGLPAQVFNITGRGSEIGDYVEHEEVNFINTGSTPVGORIGK 240
QY 241 LAQMRPIMELGSKDAGIVLADADLNNAKQIVAGAYDSGQCTAIKRYLVVEEVADEL 300
DB 241 LAQMRPIMELGSKDAGIVLADADLNNAKQIVAGAYDSGQCTAIKRYLVVEEVADEL 300
QY 301 AEKISENVAKLSVGDPPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRRDRLLT 360
DB 301 AEKISENVAKLSVGDPPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRRDRLLT 360
QY 361 PGHFDHTVLDKLAWEFPFPIIPRIKYDAEBAVAIANKSDFGLOSSVFTTRDFOKAFDI 420
DB 361 PGHFDHTVLDKLAWEFPFPIIPRIKYDAEBAVAIANKSDFGLOSSVFTTRDFOKAFDI 420
QY 421 ANKLEVGTVHINKKTGRGPDNPFPLGLKSGAGVQGIKRSIEAMTVKSIIVLDMK 475
DB 421 ANKLEVGTVHINKKTGRGPDNPFPLGLKSGAGVQGIKRSIEAMTVKSIIVLDMK 475

RESULT 4
Q97QTO :
AC Q97QTO PRELIMINARY; PRT; 474 AA.
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent.
GN OrderedLocustNames=SP1119;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1313;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; Pubmed=11463916; DOI=10.1126/science.1061217;
Tetcelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,

RA Peterson S.N., Heidelberg J.F., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzaple E.K., Knout H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT Complete genome sequence of a virulent isolate of *Streptococcus*
RT pneumoniae.";
RL Science 293:498-506 (2001).
DR EMBL; AE007413; AAK75230.1; -.
DR PIR; E95129; B95129.
DR HSSP; Q59931; IEUH.
DR TIGR; SP1119; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR Complete proteome.
DR KW complete proteome.
SQ SEQUENCE 474 AA; 51127 MW; 89D8C05F02AC0048 CRC64;

Query Match 75.7%; Score 1812; DB 2; Length 474;
Best Local Similarity 74.1%; Pred. No. 2.4e-100; Indels 0; Gaps 0;
Matches 349; Conservative 57; Mismatches 65; Indels 0; Gaps 0;

QY 5 YQNYVNGEKSSVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALTYVER 64
DB 4 YQNYVNGEKSSVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAAALTYVER 63
QY 65 AOTLHKADIIERDKKEIATVLAKEISKAYNASVTEVATADLIRYAAEEGIRLSTAD 124
DB 65 AOTLHKADIIERDKKEIATVLAKEISKAYNASVTEVATADLIRYAAEEGIRLSTAD 123
QY 125 GGMOMASSTGHKLAVIRORPGIVLAIPAPYNPVNLGSKIAPALIGGNVMPKPPTQGSV 184
DB 124 GGMOMASSTGHKLAVIRORPGIVLAIPAPYNPVNLGSKIAPALIGGNVMPKPPTQGSV 183
QY 185 SGLVAKAEAGLPAQVFNITGRGSEIGDYVEHEEVNFINTGSTPVGORIGLQAM 244
DB 184 SGLVAKAEAGLPAQVFNITGRGSEIGDYVEHEEVNFINTGSTPVGORIGLQAM 243
QY 245 RPIMLELGGKDGIVLADADLNNAKQIVAGAYDSGQCTAIKRYLVVEEVADELAETI 304
DB 245 RPIMLELGGKDGIVLADADLNNAKQIVAGAYDSGQCTAIKRYLVVEEVADELAETI 303
QY 305 SENVAKLSVGDPPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRRDRLT 364
DB 304 SENVAKLSVGDPPDNATVPTVIDNSADFIESIYVDARQGAKEINEFRRDRLT 363
QY 365 DHVTLDMKLAWEFPFPIIPRIKYDAEBAVAIANKSDFGLOSSVFTTRDFOKAFDIANKL 424
DB 364 DHVTLDMKLAWEFPFPIIPRIKYDAEBAVAIANKSDFGLOSSVFTTRDFOKAFDIANKL 423
QY 425 EVGTVHINKKTGRGPDNPFPLGLKSGAGVQGIKRSIEAMTVKSIIVLDMK 475
DB 424 EVGTVHINKKTGRGPDNPFPLGLKSGAGVQGIKRSIEAMTVKSIIVLDMK 474

RESULT 5
Q8DP57 :
AC Q8DP57 PRELIMINARY; PRT; 474 AA.
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC
DE 1.2.1.9).
GN Name=gadN; OrderedLocustNames=SP11028;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

OK NCBI_TaxID=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 RA DOI=10.1128/JB.183.19.5709-5717.2001;
 RA Hoskins J., Alborn W.E. Jr., Arnold U., Blaszcak L.C., Burgett S.,
 RA Dehoff B.S., Battem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz B.J., Lu J., Matsushita P.,
 RA McLaren S.M., McHenry M., Mclester K., Mundy C.W., Nicas T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun P.-M., Winkler M.B., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Kosteck P.R. Jr., Skatrud P.L.,
 RA Glass J.I.;
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL: AE008475; AAK99832.1; -.
 DR PIR: D98000; D98000.
 DR HSSP: Q5931; 1E0H.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyde_dehydratog.
 DR Pfam: PF00171; Aldehyd_1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW COMPLETE proteome.
 SQ SEQUENCE 474 AA; 51076 MW; 87BC0A8BD75E21 CRC64;

Query Match 75.3%; Score 1802; DB 2; Length 474;
 Best Local Similarity 73.7%; Pred. No. 9.5e-100;

Matches 347; Conservative 58; Mismatches 66; Indels 0; Gaps 0;

QY 5 YONTYNGEWSVNOQIEILSPIDSSLGFPVPMTRREVDHAMKAGREATPAMALTVYR 64
 DB 4 YONLVNGKSSSQEQLTITSPINOELGVPMTOFEDEANQAAALPAMRALSAIR 63
 QY 65 AQLYKADIIERDKKEIATVLAKEISKAYNSVTEVETADLIRYAAEGRLSTADE 124
 DB 64 AAVLHKTAAILERDKKEIGITLAKVAKGKKAIGEVETADLIRYAAEGRLSTADE 123
 QY 125 GGMDSSTGHKLVIRROVGVLAIPVNPVNLGSKTAPRLIGANNVMPFPQGSV 184
 DB 124 GGFPEASAKKLAVRREBPVGLAIPVNPVNLGSKTAPRLIGANNVMPFPQGSV 183
 QY 185 SGLVLAKEAEAGLPAGVFNITGRGSEIGDYIVEHEVNFIFGTSTPVGRIGLAGM 244
 DB 184 SGLLAKAFEBEAGIPAGVFNITGRGSEIGDYIIEKVNFINFTGSTPIGERIGLAGM 243
 QY 245 RPTMELGSKDAGIVLADLDNAKQIVAGAYDSGRCCTAIKRVLVVEVADELAEKI 304
 DB 244 RPTMELGSKDALVLEDADLEHAQIVAGAFSYSGRCCTAIKRVLVESVADKLATLL 303
 QY 305 SERVVALSTGDDPPDNATVTPVINDSNADPTESLVNAROKAKELNEFKDGLLTPGLF 364
 DB 304 QEVSKLTATGDDPPDNADITPVINDSNADPTFWGLIEAQAQEAALPIRKEGLLMPVLF 363
 QY 365 DHVTLDMKLAEEPPFPILPIIRVKDAEVAIVANKSDGLGSSVTRPFOKAFDIANL 424
 DB 364 DQVTKMKVAMEPPFPVPIIRVASVEEIALIAPANSSEBGLGSSVTTNDPKAFELAEKL 423
 QY 425 EVGTVAHINKTGKGPDPNFPFLGLKSGAGVQIRYSIEANTVKSIVLDMK 475
 DB 424 EVGTVAHINKTGKGPDPNFPFLGLKSGAGVQIRYSIEANTVKSIVPDK 474

RESULT 6
 ID GAPPN STRMU STANDARD; PRT; 475 AA.
 AC 059931;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9)

DE (Non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase)
 DE (glyceraldehyde-3-phosphate dehydrogenase [NADP+]) (triosephosphate
 DE dehydrogenase).
 GN Name=gapn; OrderedLocustNames=SMU.676;
 OS Streptococcus mutans; Lactobacillales; Streptococcaceae;
 OC Bacteria; Firmicutes; Streptococcus.
 OC Streptococcus.
 NC NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NGS / Serotype C;
 RA MEDLINE=95270576; PubMed=7751269;
 RA Boyd D.A., Cviltkovitch D.G., Hamilton I.R.;
 RT "Sequence, expression, and function of the gene for the
 RT nonphosphorylating, NADP-dependent glyceraldehyde-3-phosphate
 RT dehydrogenase of Streptococcus mutans.";
 RL J. Bacteriol. 177:2622-2627(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -I- FUNCTION: Important as a means of generating NADPH for
 CC biosynthetic reactions.
 CC -I- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + NADP(+) + H(2)O
 CC = 3-phospho-D-glycerate + NADPH.
 CC -I- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L38521; AAA91091.1; -.
 CC EMBL: AE014911; AAN58410.1; -.
 DR PIR: AS7151; A57151.
 DR PDB: 1EHU; X-ray; A/B/C/D=1-475.
 DR PDB: 1Q11; X-ray; A/B/C/D=1-475.
 DR PDB: 1Q16; X-ray; A/B/C/D=1-475.
 DR PDB: 2E0H; X-ray; A/B/C/D=1-475.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR Pfam: PF00171; Aldehyd_1.
 DR TIGRFam: TIGR01804; BADH; 1.
 DR TIGRFam: TIGR01722; MMSDH; 1.
 DR TIGRFam: TIGR01780; SSADH; 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; FALSE NEG.
 KW 3D-structure; Complete proteome; NADP; Oxidoreductase.
 KW NP BIND 230 235
 FT ACT_SITE 250 250
 FT ACT_SITE 284 284
 FT CONFLICT 58 58
 FT CONFLICT 85 85
 FT CONFLICT 347 347
 FT STRAND 4 4
 FT STRAND 7 9
 FT TURN 10 11
 FT STRAND 12 14
 FT STRAND 19 23
 FT TURN 25 27
 FT STRAND 30 35
 FT HELIX 39 58
 FT HELIX 61 77
 FT TURN 78 78

FT HELIX 79 90
FT TURN 91 91
FT HELIX 94 115
FT TURN 116 117
FT STRAND 120 124
FT HELIX 125 127
FT TURN 128 128
FT HELIX 130 132
FT TURN 133 134
FT STRAND 135 143
FT STRAND 146 150
FT TURN 153 154
FT TURN 156 157
FT HELIX 158 169
FT TURN 170 171
FT STRAND 173 177
FT TURN 181 182
FT HELIX 183 196
FT TURN 197 197
FT TURN 200 202
FT STRAND 203 205
FT HELIX 210 219
FT TURN 221 222
FT STRAND 225 229
FT HELIX 232 241
FT TURN 242 244
FT STRAND 247 250
FT STRAND 255 259
FT TURN 261 262
FT HELIX 265 277
FT HELIX 278 281
FT STRAND 287 293
FT HELIX 294 309
FT TURN 310 310
FT STRAND 313 313
FT HELIX 316 318
FT TURN 319 319
FT STRAND 322 322
FT HELIX 328 343
FT TURN 344 345
FT STRAND 347 348
FT TURN 354 355
FT TURN 356 357
FT STRAND 358 359
FT STRAND 362 365
FT TURN 369 370
FT HELIX 372 374
FT TURN 375 375
FT STRAND 382 388
FT HELIX 391 400
FT STRAND 403 404
FT STRAND 406 410
FT HELIX 414 423
FT STRAND 428 431
FT TURN 440 441
FT STRAND 444 444
FT STRAND 447 448
FT TURN 449 450
FT STRAND 451 452
FT STRAND 455 455
FT HELIX 457 463
FT TURN 464 464
FT STRAND 465 473
SQ SEQUENCE 475 AA; 51194 MW; FOA2770AB9552DC CRC64;

Query Match 71.0%; Score 1700; DB 1; Length 475;
Best Local Similarity 66.9%; Pred. No. 1.2e-93;
Matches 318; Conservative 75; Mismatches 82; Indels 0; Gaps 0;

Qy 1 LTRKYONYVNGEMKSSVNOIEILSPIDSSLGFPVPMATREEDVHAMKAGREALPMAAALT 60
Db 1 MTKQYKNVYVNGEMKLSNEIKIYEPASGALGVPAMSTEEVDVYVASAKKQAPARSL 60

Qy 61 VYERAOYLHKADIIERDKKEIATVLAKEISKAYNASVTEVVRTADLIRYAAEGIRLST 120
Db 61 YIERAAVYIHKVADIIIRDKKEIGAVLSKEVAGYSAVSEVVRTAIIINYAAEGIRMEG 120
Qy 121 SADEGKMDASTGHTKLAVIRQPVGIVLAIAEPYVFNISGSKIAIPALIGVYVMPKPT 180
Db 121 EVLEGGSPFAASKKIAYVRRREPVGIVLAISFPNPNVLAGSKIAIPALVAGVIAKPT 180
Qy 181 QGSVSGVLIAKAPAEKGLPAGVFNITITGRGSSIGDIYVEHEEVFNITGSPVPGRIK 240
Db 181 QGISGGLIAEAPAEKGLPAGVFNITITGRGSSIGDIYVHQVNFNITGSGIGRIK 240
Qy 241 LAGMRPIMELGKAGITLADADLDMAKQIVAGAYDVGSGRCIAIKRVVVEEPADEL 300
Db 241 MAGMRPIMELGKAGITLADADLDMAKQIVAGAYDVGSGRCIAIKRVVVEEPADEL 300
Qy 301 AEKISENVAKLSVDPEDNATVTPVDDSDADPISLVDPAKOKAKEINPEADGRLLT 360
Db 301 VEKIREKVLALITGNPEDDADITPLIDTSADYVEGLINDADKGAALTEIKREGNLIC 360
Qy 361 PGLFDHVTLDMLAMEPPGPILPIIRYDAEBAVAIAKSDPGLQSVFTRDPOKAPDI 420
Db 361 PILFDKVTIDMLAMEPPGPVLPPIIRYTSVEEALIESKNSEYGLQASIFNDPFAFGI 420
Qy 421 ANKLEVGTVHINKTGRGPDNPPFLGKSGAGVQGIRISITAMTVKSIIVLDMK 475
Db 421 AEQLEVGTVHINKTGRGPDNPPFLGKSGAGVQGIRISITAMTVKSVVDIK 475

RESULT 7
Q8POF5 PRELIMINARY; PRT; 475 AA.
ID Q8POF5
AC Q8POF5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.
GN Name=gadN; Ordered locus names=spym18_1983;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AE010058; AAL97978.1; -.
DR HSSP; Q59931; 1E0H.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydrog.
DR Pfam; PF00171; Aldehyd_1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR Complete proteome.
SQ SEQUENCE 475 AA; 50381 MW; 335E29B77212DFC5 CRC64;

Query Match 68.7%; Score 1645; DB 2; Length 475;
Best Local Similarity 65.9%; Pred. No. 2.4e-90;
Matches 313; Conservative 75; Mismatches 87; Indels 0; Gaps 0;

Qy 1 LTRKYONYVNGEMKSSVNOIEILSPIDSSLGFPVPMATREEDVHAMKAGREALPMAAALT 60
Db 1 MTKQYKNVYVNGEMKLSNEIKIYEPASGALGVPAMSTEEVDVYVASAKKQAPARSL 60
Qy 61 VYERAOYLHKADIIERDKKEIATVLAKEISKAYNASVTEVVRTADLIRYAAEGIRLST 120

```
Db 61 YVERAAVTHKADILVRDAEKIGALISKEVAKHKAABSEVIRTAETIINYAAEEGRLMEG 120
121 SADEGKMDASTGHTLAIVRRPVGIVLAIAPYNYVNI:SGSKIAIPALIGGVVMEKPEPT 180
122 EYLEGGSPFAASKKKAIVRRBPVGIVLAI:SPFNYPVNLASKIAIPALIGGVVALKPEPT 180
Db 121 EYLEGGSPFAASKKKAIVRRBPVGIVLAI:SPFNYPVNLASKIAIPALIGGVVALKPEPT 180
Qy 181 QGSVSGVLAKAPAEAGLPAGVFNITTRGSEIGDYIVHEEYNFINFTGSTPVGORIGK 240
181 QGISGGLLAEAFABAGIPAGVFNITTRGSGVIGDYIVHEEYNFINFTGSTPVGORIGK 240
Qy 241 LAGMPRIMELGKQSAIVLEADADLAAKNIVAGFYGSGRCITAVKVLVMDVADQL 300
241 LAGMPRIMELGKQSAIVLEADADLAAKNIVAGFYGSGRCITAVKVLVMDVADQL 300
Db 301 AEKISENVAKLSVGPDPNATVTPVIDNSADFIETSLVVDAROKAKELNEFKDGRLLT 360
301 AEKITLVKLSVGMPEDDADITPLIDTSADFPVEGLIKDNTDKGATLTAFFNRGNLIS 360
Qy 361 PGLFDHVTLDMLKLAWEPPGPPLPIIRYKDAEBAVAIANKSDPGLQSSVFTTRDFOKAPDI 420
361 PVLFDHVTLDMLKLAWEPPGPPLPIIRYKDAEBAVAIANKSDPGLQSSVFTTRDFOKAPDI 420
Db 361 PVLFDHVTLDMLKLAWEPPGPPLPIIRYKDAEBAVAIANKSDPGLQSSVFTTRDFOKAPDI 420
Qy 421 ANKLEVTGTHINNKTRGPDNPPFLGLKSGAGVQGIKSYSTEAMTNVKSIVLDMK 475
421 AEOLVGVTHLNKKTORGTDNPFILGAKKSGAGVGKYSIEAMTNVKSIVVFDIQ 475
```

RESULT 8

Q99267 PRELIMINARY; PRT; 475 AA.

```
AC Q99267;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9).
GN Name=gapN; OrderedlocusNames=SPY1371;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RC MEDLINE=2119264; PubMed=11296296; DOI=10.1073/pnas.01559398;
RA Ferreira J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Szatec S., Suvorov A.N., Kenon S., Lai H.S., Lin S.P.,
RA Qian Y., Clifton S.W., Roe B.A., McLaughlin R.E.;
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
RL EMBL: AB006575; AAK34198.1; -.
DR HSPF; Q59931; IEUH.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 475 AA; 50368 MW; BC24957621AE0FC9 CRC64;
```

Query Match 68.5%; Score 1639; DB 2; Length 475;

Best Local Similarity 65.5%; Pred. No. 5.4e-90;

Matches 311; Conservative 77; Mismatches 87; Indels 0; Gaps 0;

```
Qy 1 LKEXYONYNGEKSSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAAALYT 60
1 LKEXYONYNGEKSSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAAALYT 60
Db 1 MAKQYKLVNGEKLSENEITTYAIPATGELGSPVPMTOAEVDVAVAAKALSDWRALS 60
1 MAKQYKLVNGEKLSENEITTYAIPATGELGSPVPMTOAEVDVAVAAKALSDWRALS 60
Qy 61 YVERAAYTHKADILVRDAEKIGALISKEVAKHKAABSEVIRTAETIINYAAEEGRLMEG 120
61 YVERAAYTHKADILVRDAEKIGALISKEVAKHKAABSEVIRTAETIINYAAEEGRLMEG 120
Db 61 YVERAAYTHKADILVRDAEKIGALISKEVAKHKAABSEVIRTAETIINYAAEEGRLMEG 120
```

```
Qy 121 SADEGKMDASTGHTLAIVRRPVGIVLAIAPYNYVNI:SGSKIAIPALIGGVVMEKPEPT 180
121 EYLEGGSPFAASKKKAIVRRBPVGIVLAI:SPFNYPVNLASKIAIPALIGGVVALKPEPT 180
Db 121 EYLEGGSPFAASKKKAIVRRBPVGIVLAI:SPFNYPVNLASKIAIPALIGGVVALKPEPT 180
Qy 181 QGSVSGVLAKAPAEAGLPAGVFNITTRGSEIGDYIVHEEYNFINFTGSTPVGORIGK 240
181 QGISGGLLAEAFABAGIPAGVFNITTRGSGVIGDYIVHEEYNFINFTGSTPVGORIGK 240
Qy 241 LAGMPRIMELGKQSAIVLEADADLAAKNIVAGFYGSGRCITAVKVLVMDVADQL 300
241 LAGMPRIMELGKQSAIVLEADADLAAKNIVAGFYGSGRCITAVKVLVMDVADQL 300
Db 301 AEKISENVAKLSVGPDPNATVTPVIDNSADFIETSLVVDAROKAKELNEFKDGRLLT 360
301 AEKITLVKLSVGMPEDDADITPLIDTSADFPVEGLIKDNTDKGATLTAFFNRGNLIS 360
Qy 361 PGLFDHVTLDMLKLAWEPPGPPLPIIRYKDAEBAVAIANKSDPGLQSSVFTTRDFOKAPDI 420
361 PVLFDHVTLDMLKLAWEPPGPPLPIIRYKDAEBAVAIANKSDPGLQSSVFTTRDFOKAPDI 420
Db 361 PVLFDHVTLDMLKLAWEPPGPPLPIIRYKDAEBAVAIANKSDPGLQSSVFTTRDFOKAPDI 420
Qy 421 ANKLEVTGTHINNKTRGPDNPPFLGLKSGAGVQGIKSYSTEAMTNVKSIVLDMK 475
421 AEOLVGVTHLNKKTORGTDNPFILGAKKSGAGVGKYSIEAMTNVKSIVVFDIQ 475
```

RESULT 9

Q8K707 PRELIMINARY; PRT; 467 AA.

```
AC Q8K707;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.
GN Name=gapN; OrderedlocusNames=SPY3_1045;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RC MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
RL EMBL: AB014157; AAM79652.1; -.
DR HSPF; Q59931; IEUH.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 467 AA; 49478 MW; 287AEDC69638B2EB CRC64;
```

Query Match 67.6%; Score 1618; DB 2; Length 467;

Best Local Similarity 66.0%; Pred. No. 9.5e-89;

Matches 308; Conservative 74; Mismatches 85; Indels 0; Gaps 0;

```
Qy 9 VNGEKSSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAAALTYVERAAYL 68
9 VNGEKSSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAAALTYVERAAYL 68
Db 1 NMGEXKLVSENEITTYAIPATGELGSPVPMTOAEVDVAVAAKALSDWRALSYVERAAYL 60
1 NMGEXKLVSENEITTYAIPATGELGSPVPMTOAEVDVAVAAKALSDWRALSYVERAAYL 60
Qy 69 HKAADILVRDAEKIGALISKEVAKHKAABSEVIRTAETIINYAAEEGRLMEG 128
69 HKAADILVRDAEKIGALISKEVAKHKAABSEVIRTAETIINYAAEEGRLMEG 128
Db 61 HKAADILVRDAEKIGALISKEVAKHKAABSEVIRTAETIINYAAEEGRLMEG 120
```

[illegible]

OY		123	DEGGMDASTGHKLAVIRROPVGIATAPNYPVNLSGSKTAPALIGNVVMFKPPTQG	182
D5		131	INSDNPFGSKDKDSTLVERVPGLVIATLSPPNYPNLSGSKAPALIAENSVLPKSTVG	190
OY		183	SVSGVLAKAAFAEAGLPAGVFETTINGRSEISIDYIVHEEHNFINFTGSTPVGORIGKLA	242
D6		191	AISALHLAEITNAAGLPAGVLTNTVYGKSEICDYILTHEEVNFINTSSANGKHISKLA	250
OY		243	GMRPIIMEELGKGACIIVLADDLDNAAQIYAGAAYDSQGRCTAIKRVLVEEVADELAE	302
D6		251	GMIPIVTEELGKGDAIVLEEDANLETAKSIYSAGAYSGQRCTAIVRVLMKDVELYE	310
OY		303	KISEVVALSLTGDPEDPDNAVTTPVINDDNADFLIESLVANDAROKGAELNEFKDDGLTTGC	362
D6		311	LVTKKVKSELKGNPNPDVVTTIPPLIDNKADVYQTLLIDAIEKGATLIVNKKRENIMLYPT	370
OY		363	LFHDVTLDMKIAMEEPFPPIPIIRVKOAEBAVAINANKSDPFLQSSVTFTRDFOKARDIAN	422
D6		371	LFENVTLADMRLAMEEPFPGPVLPIIRYKSMDEAIELANSSEYLGQSAVFEEMHDAFYIAN	430
OY		423	KLEVGTVAHNNKTGRGPDPNPFLGLKGSQAGVGYGIRSYSTEAMTNVKSIVLDM	474
D6		431	KLDVGTVOVNNKPKBERGPDHPFLGPLTKSSGMGTGIRYSITEAMTRHSIVLNL	482
 RESULT 11 O73CX4 PRELIMINARY; PRT; 479 AA.				
ID	O73CX4			
AC	O73CX4;			
DT	05-JUN-2004 (TREMBLrel. 27, Created)			
DT	05-JUN-2004 (TREMBLrel. 27, Last sequence update)			
DT	05-JUN-2004 (TREMBLrel. 27, Last annotation update)			
DE	Glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent (EC 1.2.1.9).			
GN	Name=gapN; OrderedLocustNames=CEOE9040;			
OS	Bacillus cereus (strain ATCC 10987).			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCB1_taxonomy=222523;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=14960714; DOI=10.1093/nar/gkh258;			
RA	Raeko D.A., Ravel T., Oekstad O.A., Helgason E., Ger R.Z., Jiang L., Shores K.A., Fouts D.E., Tourasse N.J., Anguilo S.V., Kolonay J.F., Nelson W.C., Kolstoe A.B., Fraser C.M., Read T.D.; "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic adaptations and a large plasmid related to Bacillus anthracis pxoII." ; Nucleic Acids Res. 32:977-988(2004) .			
RL	EMBL; AB017267; AAC939871.1; -.			
DR	HSSP; P56533; IAMS.			
DR	TIGER; BCB0940; -.			
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	InterPro; IPRO02086; Aldehyd_dehydrog.			
DR	Pfam; PF00171; Aldehd; 1.			
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.			
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.			
KW	Complete proteome.			
SQ	SEQUENCE 479 AA; 51798 MW; E294F44609BCA580 CRC64;			
 Query Match 56.2%; Score 1345; DB 2; Length 479; Best Local Similarity 56.4%; Pred. No.2.le-72; Matches 266; Conservative 77; Mismatches 127; Indels 2; Gaps 1				
OY		5	YONTYNGEWM--SSVNOIEILSPIDDSSLGVPANTREBEVDHAMAGREALPAMAALTIVY	62
D6		7	YKFYLINGWRESSSGCETIEIPSPYLHEVIGVOAITRGDEVDEAIASAKAQAQMSBASIQ	66
OY		63	ERAQYIHKAADIITERDKKEINTVLAKESIKYMASVTSVVRPADLIRLYAAEGISTLSIA	122
D6		67	DRARYLYKMADELVNMDOEIDLIMKEVKGKQAKKKVVRPADIRITYIEBALMHGES	126
OY		123	DEGGMDASTGHKLAVIRROPVGIATAPNYPVNLSGSKTAPALIGNVVMFKPPTQG	182

```
Db 127 MMGDSFPGGTGSKLAIIORAPLGVLAIPFNNPVNLSAKLAPALIMGNNAVIFKPAATG 186
Qy 183 SVSGVLAKAFABAGLPAGVFNITTCRGSEIGDYIVHEEVNFINTGSTPVGORIGKLA 242
Db 187 AISGIKMVEALHAKGLPKGLVNVATGRSGVIGDYLVHEGIMNVSFTGTGNTGKILAKKA 246
Qy 243 GMRPIMLELGGKADGIVLADADLDMNAKOIYAGAYDSGORTAIKRYLVVBEVADLAE 302
Db 247 AMIPLVLELGGKDPGIVREDADLODPAANHIVSGAFSYSGORTAIKRYLVHENVADLVG 306
Qy 303 KISENVAKISVGDPEPDNATVTPVIDNSADFLIESLVNAROKAKELNEFPKDGRLTPG 362
Db 307 LIQEQVAKISVSGSPQDSTIVPLIDDKSADFOGLVDVAVEKATIVIGNKERNLIPT 366
Qy 363 LFDHVTLMKLAWEFPFPIIPRIKYDAEVAVALANKSDPGLQSSVFTTRDFOKAFDIAN 422
Db 367 LIDHVTEDMKVAMEEFPFPIIPRIKVSDEQALBIANKSEFGLQASVFTKDIINKAFALAN 426
Qy 423 KLEVGTVHINNKTGRGPNPFLGKSGAGVQGIKRYIEAMTNVKSIVLDM 474
Db 427 KIETGVOINGRTERGPDHPFPIGVKSGMGAGIKRKSLESMTREKVTYVNL 478
```

```
RESULT 12
O6HMX0
ID AC PRELIMINARY; PRT; 479 AA.
AC O6HMX0;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE Glyceraldehyde-3-phosphate dehydrogenase (NADP+) (EC 1.2.1.9).
GN Name=gapN; OrderedLocustNames=Bt9727.0757;
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_Taxid=180856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Rice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27."
RT Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017355; AAT59146.1; -.
DR GO; GO:0008866; F:glyceraldehyde-3-phosphate dehydrogenase (N. .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd. 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
KW Complete proteome.
SQ SEQUENCE 479 AA; 51858 MW; 679787AF11DBB8F7 CRC64;
```

Query Match 56.2%; Score 1344; DB 2; Length 479;
Best Local Similarity 56.1%; Pred No. 2.4e-72;
Matches 265; Conservative 78; Mismatches 127; Indels 2; Gaps 1;

```
Qy 5 YONTYNGEMK--SSVNOIEILSPIDDSLGFVPATREBEVDHMKAGREALPAMALTVY 62
Db 7 YKFLYNGEMRESSGSETIEIPSPYLHEVIGVQVATIRGEVDEALSAKEAQSMVAESLQ 66
Qy 63 ERAQYLHKAADIIERDKKEIATVLAKEISKAYNASVTEVTRADLIRYAEEGIRLSTSA 122
Db 67 DRAKLYKWADELVMQDEIDAIIMKEVGKGYKDAKKEVTRADFIIRYTIIEALHMHGS 126
Qy 123 DEGGKMDASTGHKLAIVIRROPVGIYLAIPVNPVNLSSGKIAPALIGNVVMFKPPTG 182
Db 127 MMGDSFPGGTGSKLAIIORAPLGVLAIPFNNPVNLAAKLAPALIMGNNAVIFKPAATG 186
Qy 183 SVSGVLAKAFABAGLPAGVFNITTCRGSEIGDYIVHEEVNFINTGSTPVGORIGKLA 242
```

```
Db 187 AISGIKMVEALHAKGLPKGLVNVATGRSGVIGDYLVHEGIMNVSFTGTGNTGKILAKKA 246
Qy 243 GMRPIMLELGGKADGIVLADADLDMNAKOIYAGAYDSGORTAIKRYLVVBEVADLAE 302
Db 247 SMIPVLELGGKDPGIVREDADLODPAANHIVSGAFSYSGORTAIKRYLVHENVADLVG 306
Qy 303 KISENVAKISVGDPEPDNATVTPVIDNSADFLIESLVNAROKAKELNEFPKDGRLTPG 362
Db 307 LYKEQVAKISVSGSPQDSTIVPLIDDKSADFOGLVDVAVEKATIVIGNKERNLIPT 366
Qy 363 LFDHVTLMKLAWEFPFPIIPRIKYDAEVAVALANKSDPGLQSSVFTTRDFOKAFDIAN 422
Db 367 LIDHVTEDMKVAMEEFPFPIIPRIKVSDEQALBIANKSEFGLQASVFTKDIINKAFALAN 426
Qy 423 KLEVGTVHINNKTGRGPNPFLGKSGAGVQGIKRYIEAMTNVKSIVLDM 474
Db 427 KIETGVOINGRTERGPDHPFPIGVKSGMGAGIKRKSLESMTREKVTYVNL 478
```

```
RESULT 13
O63FP8
ID AC PRELIMINARY; PRT; 479 AA.
AC O63FP8;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE Glyceraldehyde-3-phosphate dehydrogenase (NADP+) (EC 1.2.1.9).
GN Name=gapN; ORFNames=BtZK0750;
OS Bacillus cereus ZK.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_Taxid=288681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZK;
RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Rice H.;
RT "Complete genome sequence of Bacillus cereus ZK."
RT Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000001; AAU19493.1; -.
KW Oxidoreductase.
SQ SEQUENCE 479 AA; 51828 MW; DA9209D3A44FB04B CRC64;
```

Query Match 56.0%; Score 1341; DB 2; Length 479;
Best Local Similarity 56.1%; Pred No. 3.7e-72;
Matches 265; Conservative 78; Mismatches 127; Indels 2; Gaps 1;

```
Qy 5 YONTYNGEMK--SSVNOIEILSPIDDSLGFVPATREBEVDHMKAGREALPAMALTVY 62
Db 7 YKFLYNGEMRESSGSETIEIPSPYLHEVIGVQVATIRGEVDEALSAKEAQSMVAESLQ 66
Qy 63 ERAQYLHKAADIIERDKKEIATVLAKEISKAYNASVTEVTRADLIRYAEEGIRLSTSA 122
Db 67 DRAKLYKWADELVMQDEIDAIIMKEVGKGYKDAKKEVTRADFIIRYTIIEALHMHGS 126
Qy 123 DEGGKMDASTGHKLAIVIRROPVGIYLAIPVNPVNLSSGKIAPALIGNVVMFKPPTG 182
Db 127 MMGDSFPGGTGSKLAIIORAPLGVLAIPFNNPVNLAAKLAPALIMGNNAVIFKPAATG 186
Qy 183 SVSGVLAKAFABAGLPAGVFNITTCRGSEIGDYIVHEEVNFINTGSTPVGORIGKLA 242
Db 187 AISGIKMVEALHAKGLPKGLVNVATGRSGVIGDYLVHEGIMNVSFTGTGNTGKILAKKA 246
Qy 243 GMRPIMLELGGKADGIVLADADLDMNAKOIYAGAYDSGORTAIKRYLVVBEVADLAE 302
Db 247 SMIPVLELGGKDPGIVREDADLODPAANHIVSGAFSYSGORTAIKRYLVHENVADLVG 306
Qy 303 KISENVAKISVGDPEPDNATVTPVIDNSADFLIESLVNAROKAKELNEFPKDGRLTPG 362
Db 307 LYKEQVAKISVSGSPQDSTIVPLIDDKSADFOGLVDVAVEKATIVIGNKERNLIPT 366
Qy 363 LFDHVTLMKLAWEFPFPIIPRIKYDAEVAVALANKSDPGLQSSVFTTRDFOKAFDIAN 422
```

```
D6      367 LIDHTEEMKAMEEPFGIILPIIRVSSDEQALIELANKSEBFLQSVPFKIOKNKAFFAIAN   426
OY      423 KLEVGTVHIINKTGGRGPDNPFPLGLKSAGAVOGIRYSIEAANTNKSYIVLDM   474
Db       427 KIETGSVOINGRTERGDPHFPPFIIGVKGSGMGQGIRKISLESMTREKVTLNL   478

RESULT 14
08JUL6    PRELIMINARY;          PRT;         479 AA.
ID J08JUL6
AC Q8JUL6; O6J2U6; O6KW3;
DT 01-JUN-2003 (TREMBREL. 24, Created)
DT 21-JUN-2003 (TREMBREL. 24, Last sequence update)
DT 25-OCT-2004 (TREMBREL. 28, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent.
GN Name=gspN; OrderedLocustNames=BA0849, BAO808, GBA0849;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1392;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Reed T.D., Peterson S.N., Tourasse N.J., Baillye L.W., Paulsen I.T.,
RA Nelson K.E., Tetcelin H., Fouts D.E., Eisen U.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Ristone V., Wu M.,
RA Kolonyay J.F., Benam M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA Deboy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benson J.L., Mahmoud Y., Jiang L., Hance I.R., Weidman J.P.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Merlan W.C.,
RA Hazen A., Clime R.T., Redmond C., Thwaites J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstro A.-B., Fraser C.M.;
RA "The genome sequence of Bacillus anthracis Ames and comparison to
RL closely related bacteria.";
RL Nature 423:81-86(2003).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Rayvel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzman S.L.,
RA Fraser C.M.;
RA "Bacillus anthracis comparative genomics." ;
RL submitted (MAY-2004) to the EMBL/genbank/DDbj databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=Steiner;
RA Bretzin T.S., Bruce D., Chalacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Rice H.;
RL Submitted (JUN-2004) to the EMBL/genbank/DDbj databases.
DR EMBL; AE017026; AAP24851.1; -
DR EMBL; AE017334; AAT99962.1; -
DR EMBL; AE017225; AAT5135.1; -.
DR HSPB; Q59931; IEUH.
DR TIGR; BA0849; -.
DR TIGR; GBA0849; -.
DR GO; GO:0016491; F:oxygen reductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPRO02086; Aldehyd_dehydrg.
DR Pfam; PF00171; Aldedh; 1.
DR PROSITE; PS00070; ALDEHYDE DEHYD CYS; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KM Complete proteome.
SQ     complete protein.           51810 MW; 587029DB2C8CF652 CRC64;
```

Db	7	IKFYINGMWRSSSEETIEIEBPYLHEVIGVQVQALITRGVEADAIASAKAOKSMASLQ	66			
Qy	63	ERAQYIAKRAADIIERDKKEIATVLAKEISKAYNASVTEVTRADILRYAAEGIRLSTSA	122			
Db	67	DRAKLYIKMADEBLVMQDEIADIIMKEVGKGYDKAKKEVTRADIFIRYITEALHMGHS	126			
Qy	123	DEGGMDASSTGHKLAVIRKOPRIGVILTAIPYRYPNVLSGSKIAFPALIGNVMPKRPPTG	182			
Db	127	MMGSPFPGTISKKALIIORAPLGVVLAIAFPYYPNLSAKKALAPALIMGNVAVIFKPATGQ	186			
Qy	183	SVSGVLAKAFPAEAGLPAGVFNTTIGRSGSEIDYIVHEHNVFINFTGSTPVGORIGKA	242			
Db	187	AISGIKMVEALHKAGLPQGLVNVATGRSGVIEDYLVEHGBIMWSTGGTNGKHLAKQA	246			
Qy	243	GMRPIMELEGGKDAQIVLADADLDNNAKQIVAGAYDSGORCTAIKRVLVVEVADLAE	302			
Db	247	SMIPFLVELGGKDPETIVREDADLDQANHHIYSGASYSQRCCTAIKRVLVHENVADLVS	306			
Qy	303	KISENVAKLSVGDPDPDNAVITVPVLDNNSADPIESLIVDAROKGAKELNEFKDGRLLTGG	362			
Db	307	LVKEQVAKLSVSGPQSDSTIVPLIDDKSADVLQGVLDVAVEKGAIVIONKEERNIYPT	366			
Qy	363	LEDHTLMLKLAWEPPRPILPIIRIVKQAEVLAITANSDFGLOSVFTRDQKADIAN	422			
Db	367	LIDHTEEMKVAWEPPRPILPIIRIVSSDEQALIEIANSSEFQLQASVFPRDKINKAFIAN	426			
Qy	423	KLEVGVTHINNKTGRGPNPFLGKSGGAGVQGIIRYSIEANTVYSIYLDM	474			
Db	427	KIEFGSVQINGRTERTGRPHFPPIGVKSGMGAGIIRKLSIESMTRKRYVYLN	478			
RESULT 15						
Q72226	ID	PRELIMINARY:	PRT: 477 AA.			
AC	Q722Z6					
DT	05-JUL-2004	(TREMBLrel. 27, Created)				
DT	05-JUL-2004	(TREMBLrel. 27, Last sequence update)				
DT	05-JUL-2004	(TREMBLrel. 27, Last annotation update)				
DE	Glyceroldehyde-2-phosphate dehydrogenase, NADP-dependent.					
GN	Ordered locus names=BCE4521.					
OS	Bacillus cereus (strain ATCC 10987).					
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.					
OX	NCBI_TaxID=22523;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	PubMed=14960714; DOI=10.1093/nar/gkh258;					
RA	Raako D.A., Ravel J., Oekstad O.A., Helgeson E., Cer R.Z., Jiang L.,					
RA	Shores K.A., Fouts D.E., Tourasse N.J., Angluoli S.V., Koltony J.F.,					
RA	Neison W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;					
RT	"The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic					
RT	adaptations and a large plasmid related to Bacillus anthracis pXO1."					
RL	Nucleic Acids Res. 32:977-988(2004).					
DR	EMBL; AE017278; AA843422.1; ..					
DR	HSSP; P56533; 1A4S.					
DR	TIGR; BCE4521; ..					
DR	GO: GO:0016491; F:oxidoreductase activity; IEA.					
DR	GO: GO:0008152; P:metabolism; IEA.					
DR	InterPro; IPR002086; Aldehyd. dehydrog.					
DR	Pfam; PF00171; Aldehd., 1.					
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.					
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.					
KW	Complete proteome.					
SC	SEQUENCE 477 AA; 51638 MW; 0931DB587AE84536 CRC64;					
Query Match 55.8%; Score 1335; DB 2; Length 477;						
Best Local Similarity 55.8%; Pred. No. 8.3e-72;						
Matches 264; Conservative 78; Mismatches 129; Indels 2; Gaps 1						
Qy	5	YONYVNGWK--SSYNOIEIISPTIDSSLSGFVPAWTRFVHAMKAGREALPAMALATY	62			
Db	5	YKFTYINGMWRSSSGETIEILSPYIAHVIGVQVQALITRGVEADAIASAKAOKSMASLQ	64			
Qy	63	ERAQYIAKRAADIIERDKKEIATVLAKEISKAYNASVTEVTRADILRYAAEGIRLSTSA	122			

```
Db      :|||:| : :||| :| : ||||| ||| :| :  
65 DRAKLYKMADELVNMQDEIADIIMKEVKGKADAKEVETADFIIRYIEBALHMGES 124  
Qy      123 DEGGKMDASTGHKLAIVIRQPGVIVLAIPYNPVNLGGSKTAPALIGNVMMFKPPTQG 182  
Db      125 MMGDSFPGGTGSKLAIIQAPLGVVLAIPFNYPVNLGAAPALIGNAVIFKPAIQTG 184  
Qy      183 SVSGVLAKAPAEAGLPAGVFNITIGRSGEIGDYIVHEBEVNFINTGSPVQGRIGKLA 242  
Db      185 AISGIKMVEALHKAGLPKGLVNVATGRSGVIGDYIVHEBGIDMVSTGTGNTGKHLAKKA 244  
Qy      243 GKRPMLEIGGKADAGIVLADLDNNAKOIVAGADVSGORCTAIKRVLVVEBVADELAE 302  
Db      245 AMITVLIELGKDPGIVREDADLDQAAHIVSGAFSGORCTAIKRVLVHENVADELVS 304  
Qy      303 KISENVAKLSVGDPPDNATVTPVIDNSADPTIESIVVDAROKAKEINEFKRDGLTPG 362  
Db      305 LIQOVAKLSVSGPEQDSTIVPLIDKSDAPVOGLVDDAVEKGATTVIGNKRERLIPT 364  
Qy      363 LEDHTLDMKLAMEBEPFGLPIIRVKADEBAVAIANKSDPGLQSVFTPDFOKAFDIAN 422  
Db      365 LIDHTVEDMTVAMEBEPFGLPIIRVSDQAIETANKSEFGLQSVFTKDIINKAFAIAN 424  
Qy      423 KLEVGTVHINKTGGPDPNPFPLGLKSGAGVQGRISIEAMTNVKSIVLDMK 475  
Db      425 KIETGSVOINGRTERGPDHPFPIGVKSGMGAQGRKSLSEMTREKVTIVNL 477
```

RESULT 16

```
Q81HE6 PRELIMINARY; PRT; 479 AA.  
ID Q81HE6  
AC Q81HE6;  
DT 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9).  
GN OrderedLocustNames=BC0868;  
OS Bacillus cereus (strain ATCC 14579 / DSM 31).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI_TaxID=226900;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;  
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,  
RA Kapetral V., Bhattacharya A., Reznik G., Mikhailova N., Lapins A.,  
RA Chu L., Marcu M., Goltsman E., Larsen N., P'Souza M., Malinas T.,  
RA Greckin T., Pusch G., Haselkorn R., Fomstein M., Ehrlich S.D.,  
RA Overbeek R., Kytrides N.C.;  
RT "Genome sequence of Bacillus cereus and comparative analysis with  
RT Bacillus anthracis.";  
RT Nature 423:87-91(2003).  
RL EMBL; AF017000; AAP0785.1; -.  
DR HSP; Q59931; 1BUH.  
DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR002086; Aldehyd_dehydrog.  
DR Pfam; PF00171; Aldehyd; 1.  
DR PROSITE; PS006070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.  
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.  
KW Complete proteome; Oxidoreductase.  
SQ SEQUENCE 479 AA; 51801 MW; F4E446B91FF8AC34 CRC64;
```

Query Match 55.7%; Score 1333; DB 2; Length 479;

Best Local Similarity 55.7%; Pred. No. 1.1e-71; Matches 263; Conservative 79; Mismatches 128; Indels 2; Gaps 1;

```
Qy      5 YONYVNGEMK--SSVNOIELSPIDSSLGFPVAMTRREVDHAMKAGREALPAMALTVY 62  
Db      7 YFYINGEMRESSGTEITLIPSPYLHEVIGVOAIRGVEDELASAKRAQSWMASISQ 66  
Qy      63 EBAQYLHKAADIIERDKETIATVLAKEISKAVNASYTBVVRTADLIRYAAEGIRLSTSA 122  
      :|||:| : :||| :| : ||||| ||| :| :  
      :|||:| : :||| :| : ||||| ||| :| :
```

```
Db      67 DRAKLYKMADELVNMQDEIADIIMKEVKGKADAKEVETADFIIRYIEBALHMGES 126  
Qy      123 DEGGKMDASTGHKLAIVIRQPGVIVLAIPYNPVNLGGSKTAPALIGNVMMFKPPTQG 182  
Db      127 MMGDSFPGGTGSKLAIIQAPLGVVLAIPFNYPVNLGAAPALIGNAVIFKPAIQTG 186  
Qy      183 SVSGVLAKAPAEAGLPAGVFNITIGRSGEIGDYIVHEBEVNFINTGSPVQGRIGKLA 242  
Db      185 AISGIKMVEALHKAGLPKGLVNVATGRSGVIGDYIVHEBGIDMVSTGTGNTGKHLAKKA 246  
Qy      243 GKRPMLEIGGKADAGIVLADLDNNAKOIVAGADVSGORCTAIKRVLVVEBVADELAE 302  
Db      247 SMIPVLIELGKDPGIVREDADLDQAAHIVSGAFSGORCTAIKRVLVHENVADELVS 306  
Qy      303 KISENVAKLSVGDPPDNATVTPVIDNSADPTIESIVVDAROKAKEINEFKRDGLTPG 362  
Db      307 LIKAQVABLSTSGPEQDSTIVPLIDKSDAPVOGLVDDAVEKGATTVIGNKRERLIPT 366  
Qy      363 LEDHTLDMKLAMEBEPFGLPIIRVKADEBAVAIANKSDPGLQSVFTPDFOKAFDIAN 422  
Db      367 LIDHTVEDMTVAMEBEPFGLPIIRVSDQAIETANKSEFGLQSVFTKDIINKAFAIAN 426  
Qy      423 KLEVGTVHINKTGGPDPNPFPLGLKSGAGVQGRISIEAMTNVKSIVLDM 474  
Db      427 KIETGSVOINGRTERGPDHPFPIGVKSGMGAQGRKSLSEMTREKVTIVNL 478
```

RESULT 17

```
Q8XHP4 PRELIMINARY; PRT; 482 AA.  
ID Q8XHP4  
AC Q8XHP4;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.  
GN Name=gapN; OrderedLocustNames=CPE2438;  
OS Clostridium perfringens.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI_TaxID=1502;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=13;  
RX MEDLINE=21664373; PubMed=11792842; DOI=10.1073/pnas.022493799;  
RA Shintzu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,  
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;  
RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
RT flesh-eater.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).  
RL EMBL; AP003194; BAB82144.1; -.  
DR HSP; Q59931; 1BUH.  
DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR Pfam; PF00171; Aldehyd; 1.  
DR PROSITE; PS006070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.  
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.  
KW Complete proteome.  
SQ SEQUENCE 482 AA; 52763 MW; 2DD7DA4A3BC03EC4 CRC64;
```

Query Match 55.3%; Score 1323; DB 2; Length 482;

Best Local Similarity 52.3%; Pred. No. 4.4e-71; Matches 251; Conservative 99; Mismatches 124; Indels 6; Gaps 3;

```
Qy      2 TRE-YONYVNGEM--KSSVNOIELSPIDSSLGFPVAMTRREVDHAMKAGREALPAMA 57  
Db      3 TKENIFVPMGNGFNENKSDKLDIKSPDLSLGRISQSLSEDDVDKTIIDNAEKQKSWN 62  
Qy      58 ALTVERRQYVLAHKAADIIERDKETIATVLAKEISKAVNASYTBVVRTADLIRYAAEGIR 117  
Db      63 EYPLNERHVLVYKTDLDLEHKEELANIMIEVAKDKSALSBSVRTDYIRFTADAKS 122  
Qy      118 ISTSADEGGKMDASTGHKLAIVIRQPGVIVLAIPYNPVNLGGSKTAPALIGNVMMFK 177  
      :|||:| : :||| :| : ||||| ||| :| :  
      :|||:| : :||| :| : ||||| ||| :| :
```

```

Db      123 MDGESIPDNFPGFNRSKISVTRREPLGVTLAISPFNYPVNLASAKIAPALVAGNSVTLK 182
Qy      178 PPTQGSVSGVLVAKAAREAGLPAGVNTTITGRGSEIGDYIVHEBEVNFINTGSPVQGR 237
Db      183 PATQGSISALFLAKMFHDAGLPQGVNTITVGRGSEIGDYCVTDKIDFINFSTSTEVGR 242
Qy      238 IGLTAGRPIMELGGKDGIVLADADLDNAKOIYAGAVDYGORCTAIKRVLYVEEVA 297
Db      243 ISBITWKPULMELGGDAIVLEDDLDLTNNITTAGISYISGQCTAIKRVLYVMSVA 302
Qy      298 DELAEKISENVAKLSYGDPP--DNATVTPIIDNSADFIESTLVVDARQKAKELNEFKD 355
Db      303 DKIVETLKEKISLTKNGNPLEVEDVTIIVPLIDNKADFWVELIDDAKEKATLVCGGHE 362
Qy      356 GRLLTPEGLFDVTLDMKLAMEEPPGPIPLIRVDAEENVAIANKSDPGLQSSVTRDFO 415
Db      363 GNLIETATLFDNVTITDRMLAEPEPGPVLPIIRINDDEAIKIANASYGLQSSVTRDIN 422
Qy      416 KARDIANKLEVGVTHNNKTRGPDNPPLGLKSGAGVQGIYRSIEMNTVKSIVLDMK 475
Db      423 EAFYIANKLEVGTIVHNNKTERGPDHPPLGVKSGMGTOGIRYSIEMNSRPATVITIR 482

```

RESULT 18

```

O9KAO0      PRELIMINARY;      PRT;      481 AA.

```

```

AC      O9KAO0;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE      NADP-dependent glyceralddehyde-3-phosphate dehydrogenase (EC
1.2.1.9).
GN      Name=gapN; Ordered locusNames=BH2237;
OC      Bacillus halodurans.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=86665;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C-125;
RX      MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA      Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Maesu N.,
RA      Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RT      "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT      halodurans and genomic sequence comparison with Bacillus subtilis."
RL      Nucleic Acids Res. 28:4317-4331(2000).
DR      EMBL: AP001514; BAB05956.1; -.
DR      PIR: E83929; E83929.
DR      HSSP: Q59931; IEUH.
DR      GO: GO:0016491; F:oxidoreductase activity; IEA.
DR      GO: GO:0008152; P:metabolism; IEA.
DR      InterPro: IPR002086; Aldehyd_dehydrog.
DR      Pfam: PF00171; Aldehyd. 1.
DR      PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR      PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW      Complete proteome; Oxidoreductase.
SQ      SEQUENCE 481 AA; 51452 MW; 82769AC55C95F0B9 CRC64;

```

Query Match 52.0%; Score 1244; DB 2; Length 481;

Best Local Similarity 51.4%; Pred. No. 2.3e-66;

Matches 240; Conservative 95; Mismatches 130; Indels 2; Gaps 1;

```

Qy      10 NGEWKS--VNOIETLSPIDDSIGFVPAMTREVVDHAKGREALPAMALTVYERAOY 67
Db      15 NGMWESRTGERISISAPAGVALGISIPALSGEVDADIGADAQKIRPIRHERVDL 74
Qy      68 LHAADIIERDKIEATVLAKEISKAYNSVTEVVTADIRYAAEGRILTSADGEGK 127
Db      75 LYVADILBERKIIIGELIHEVAKPKKSIIGVSRITADIRYADALKINSETTLKGDQ 134
Qy      128 MDASTGKHLAVIRROPVGIATAPYNYPVNLGSKIAPALIGNVVMPKPPQGSVSG 187
Db      135 FKSGSSKKTIALVEREPLGVTLAISPFNYPVNLAAKIAPALVIGNTVVFEDALQSGISGT 194

```

```

Qy      188 VLAKAFAEAGPAGVNTTITGRGSEIGDYIVHEBEVNFINTGSPVQGRIGLAKMRPI 247
Db      195 KVEALADAGABEGITIQVTTGRGSYIGDLVHEPGIDMTTFGGITGRISISKAMITV 254
Qy      248 MLELGKDGAGIVLADADLDNAKOIYAGAVDYGORCTAIKRVLYVEEVADEIAEKISEN 307
Db      255 VLELGKDPALVLDADDLKLTASQIVSAGFSGQRCTAIKRVFVODSVADQVANIKNL 314
Qy      308 VAKLSVGDPPDNATVTPVIDNSADFIESTLVVDARQKAKELNEFRDGRLLTPGLFDHY 367
Db      315 VEQLTFGSDEDDADITPVIDEKSAAFIQGLIDDALENGATLLSGNRQGNLSPILLDDV 374
Qy      368 TIDMKLAMEEPPGPIPLIRVDAEENVAIANKSDPGLQSSVTRDFOCAFDIANTLEFG 427
Db      375 TPAMRVAMEEPPGPIPLIRVDAEENVAISLSNOSDYGLOASIFTKOTDRAINIGKLEVG 434
Qy      428 TVHNNKTRGPDNPPLGLKSGAGVQGIYRSIEMNTVKSIVLDM 474
Db      435 TVHINAKTERGPDHPPLGVKSGGLGVQSIKESLSMTREYTVLNL 481

```

RESULT 19

```

O9SNX8      PRELIMINARY;      PRT;      496 AA.

```

```

AC      O9SNX8;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE      Nonreversible glyceralddehyde-3-phosphate dehydrogenase (EC
1.2.1.9).
OS      Apium graveolens (Celery).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC      campanulids; Apiales; Apiaceae; Apioideae; apioid superclade;
OC      Apium clade; Apium.
OX      NCBI_TaxID=4045;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=leaf;
RX      MEDLINE=20438225; PubMed=10982446; DOI=10.1104/pp.124.1.321;
RA      Gao Z., Loeschner W.H.;
RT      "NADPH supply and mannitol biosynthesis. Characterization, cloning,
RT      and regulation of the non-reversible glyceralddehyde-3-phosphate
RT      dehydrogenase in celery leaves."
RL      Plant Physiol. 124:321-330(2000).
DR      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=leaf;
RA      ZhiFang G., Loeschner W.H.;
RX      Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF196292; AAF08296.1; -.
DR      HSSP: Q59931; IEUH.
DR      GO: GO:0016491; F:oxidoreductase activity; IEA.
DR      GO: GO:0008152; P:metabolism; IEA.
DR      InterPro: IPR002086; Aldehyd_dehydrog.
DR      Pfam: PF00171; Aldehyd. 1.
DR      PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR      PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW      Oxidoreductase.
SQ      SEQUENCE 496 AA; 53173 MW; D86AB4C01E0A3642 CRC64;

```

Query Match 49.0%; Score 1171.5; DB 2; Length 496;

Best Local Similarity 50.3%; Pred. No. 5.3e-62;

Matches 240; Conservative 83; Mismatches 143; Indels 11; Gaps 5;

```

Qy      5 YQNVNNGEKK--SYVNOIETLSPIDDSIGFVPAMTREVVDHAKGREALPAMALTVY 62
Db      16 FYTISGEMKSSSGKSVAILINPTTRMTQFOACQESVNNKAMEYAKKVOQWAKTPIW 75
Qy      63 EBAOYLHKAADIIERDKIEATVLAKEISKAYNSVTEVVTADIRYAAEGRILTSSTA 122
Db      76 KBAELHLKAAALIKKHAALADCVLKEIAKPAKDSVTEVVRSGDLVSYCAEGVRL--L 132

```

Qy	123	DECKM--DASTGH---KLAIVIRROVGVILAPVPPVNLGSKKAPALIGNVMPK	177
	133	GEKFLVSDSPFQENERTKTCIUSKIPDLGYIALPPENPVNLVASKIGEPALIAQNALVLK	192
Qy	178	PLPFGVSGLVLAKAPAEAGLPAGVPPNTITGRGSEIGDYIEHEBEVNFINTGSTPVGQR	237
Db	193	PLPFGAAVACIMHMYCHTLAGFPKGLISCTIGSGSEIGDFLTMPBWMCSIFSTFG-DIGIA	251
Qy	238	IGKLAMGRPIMTELGSCKDAGIVLADADILNMAKQIIVAGAYDVSQGRCAIRFVLVEEVA	297
Db	252	ISKAGNVPLQMLGGKDACIYLEBDADLVASNTIKGFSVSGRCALIKVILWMSVA	311
Qy	298	DELAEKISENVAVGLSYGDPDPDNAVTPPVYLDNDSADPISLBYLVDAQOKAKELAFKQDR	357
Db	312	DTLVEKVNACVALLTVGPEDNSDITPVVSESSANFISELVYDAEKCATPQCOEYKEEGN	371
Qy	358	LITPGFLDHYTLDMKLAMEBFPGLIPIIRIVYDAEAAVALANKSDPFGQSSVFRDFOKA	417
Db	372	LIMPLDLIDNVKPDMRIAMEBFPGLIPVIRINSABEGHHHCANSAFGLQGVFRDINKA	431
Qy	418	FDIAKTLVEVTHVINKTGRGPNPFFGLKSGSGAGVQIGRISITAMNVKSYLDM	474
Db	432	MLISDAHSESTIIDINAPAGPHFPQGLKQSGISQGITNSIMMXTKITVYNL	488

RESULT 20			
09ZUG8			
ID	09ZUG8	PRELIMINARY;	PRT; 496 AA.
AC	09ZUG8;		
DT	01-MAY-1999 (TRMBLrel, 10, Created)		
DT	01-MAY-1999 (TRMBLrel, 10, Last sequence update)		
DT	05-JUN-2004 (TRMBLrel, 27, Last annotation update)		
DE	Purative NADP-dependent glyceralddehyde-3-phosphate dehydrogenase (Ac2g24270/pe77da.1b).		
GN	Name=Ac2g24270;		
OS	Arabiopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,		
RA	Shen M., Ronting C.M., Fraser C.M., Somerville C.R., Venter J.C.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Town C.D., Kaul S.		
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shin P.,		
RA	Tracy S.E., Banh J., Bowser L., Carninci P., Chung M.K.,		
RA	Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,		
RA	Kallin-Neumann G., Kawai U., Lam B., Lee J.M., Lin S.X.,		
RA	Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,		
RA	Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,		
RA	Toriumi M., Yamada K., Yu G., Shinzaki K., Davis R.W., Theologis A.,		
RA	Ecker J.R.;		
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RA	Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,		
RA	Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,		
RA	Deng J.M., Hayashizaki Y., Huan V.W., Lee J.M., Ishida J., Kamiya A.,		
RA	Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,		
RA	Seki M., Shin P., Tang C.C., Toriumi M., Wallender E.K., Wong C.,		
RA	Wu H.C., Yamada K., Yu G., Yian S., Shinzaki K., Ecker J.,		
RA	Theologis A., Davis R.W.;		
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.		
RN	[5]		
RP	SEQUENCE FROM N.A.		

RA Chank R., Chen H., Kim C.J., Shim N., Bowser L., Carninci P.,
 RA Chan M., Chang C.H., Dale J.M., Hayashizaki Y., Huan V.W.,
 RA Iehida J., Jones T., Kamaya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narisaka M., Nishikawa M., Onodera C.S.,
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Wong C.C., Wu H.C., Yamada K., Yu G., Yuan S.,
 RA Shinozaki K., Davis R.W., Theologis A., Eckert J.R.,
 RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC005967; AAD03388.1; -
 DR EMBL: AY037205; AAK59790.1; -
 DR EMBL: AY136409; AAM97075.1; -
 DR EMBL: BT004551; AAC04297.1; -
 DR PIR: F84634; F84634.
 DR HSP: Q59931; 1EUH.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR Interpro: IPRO02086; Aldehyd_dehydrog.
 DR Pfam: PF00171; Aldedh. 1
 DR PROSITE: PS00070; ALDEHYDE_DEHYD_CYS; UNKNOWN 1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYD_GLU; UNKNOWN 1.
 QO SEQUENCE 496 AA; 53060 MW; 9231656a5a117505 CRC64;

Query Match	48.8%;	Score 1168.5;	DB 2;	Length 496;
Best Local Similarity	50.3%;	Pred. No. 8e-62;		
Matches 240;	Conservative 84;	Mismatches 142;	Indels 11;	Gaps 5

Qy	YONVYVGEWK--SVNYQJILSPIDSSLGFPVPMTRREVDYHAKKAGBEALPMAALTYV	62
Db	16 YKTYAGGEWMTSSGSAVAKMPATRKQYKQAOCTOEAVANVELAKSOKSAKTPLM	75
Qy	63 ERAQYLHKAADIIERKEEILATLAKETSKAVMVSVEVNTADLIRYAAEEGRJLSTSA	122
Db	76 KRAELHKAALIKONKAPMAESLVEIKAPAKOSVEVEVRSGLISYCAEBGVRI---L	132
Qy	123 DECGK--DASGTH--KLAVTRPOVGLALAPYTPVNLSSSKTAPALIGENVYMF	177
Db	133 GBEKFLIJSOPGPNDRTKYCLTSKPLDVLVAIAPFPNPNVAWSKAPALIAINSIYVK	192
Qy	178 PPTQSVSGVLVAKAPAEAGLPAVFNTITGSGEISGDIYHEHEVNFINTGSPVQOR	237
Db	193 PPTQGAVSCLIHNVCHPLAGFPKGLISCTITGSGEISGDFLTMHAYANCISTGG-DTGIS	251
Qy	238 IGTLAGRPMIMELGKGAGIYLDADIDLNAKOIVAGAYDVSQORCTAIRVLVVEVA	297
Db	252 ISKKAQMIPLMELGKGACIYLDADIDLVAASNIKGGFSYSGORCTAVVAVVMESSVA	311
Qy	298 DELAEKISENVAKLSTVGDPFDNATVPYLDINSADFIESTIYVDVROKGAELNEFRKDR	357
Db	312 DELVEKVKAKVALCTVGPEENSDITVAVSSSNAFIEGLVMDKKEGATFCOEYKRBGN	371
Qy	358 LITPGLFDHVTYLDMKLAMEBPGFPLPIIRIVYDAEAAVALANKSDPFGJOSVTFPRDOKA	417
Db	372 LITPRLLDVYRPMRIAMEBPGFVVPVPLRINSVEBGINHCNMANSFGLQGVTFPKINKA	431
Qy	418 FDIANKLEVGTYHINNKTGRGDPNPFGLKGSAGVUGIRISTEAMTNVKSIVLDM	474
Db	432 IILSDMEGTVOINSAPRGDHPFPFGLKDSGIGSGVJNINSINLMTKYTYVINL	488

RESULT 21			
GAPN_MAIZE			
ID	GAPN_MAIZE	STANDARD;	PRT; 498 AA.
AC	043272;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9)		
DE	(Non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase)		
DE	(Glyceraldehyde-3-phosphate dehydrogenase [NADP+]) (triosephosphate dehydrogenase).		
GN	Name=GPN1;		
DE	Zee mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		


```

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; panicoidae; Andropogoneae; Zea.
OX NCBI_taxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. KM5330; TISSUE=shoot;
RA MEDLINE=94180387; PubMed=7545914;
RA Habentich A., Hellman U., Cerff R.;
RT "Non-phosphorylating GAPDH of higher plants is a member of the
RT aldehyde dehydrogenase superfamily with no sequence homology to
RT phosphorylating GAPDH.";
RL J. Mol. Biol. 237:165-171(1994).
CC -I- FUNCTION: Important as a means of generating NADPH for
CC biosynthetic reactions.
CC -I- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + NADP(+) + H(2)O
CC = 3-phospho-D-glycerate + NADPH.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ch/announce/
CC or send an email to license@ebi-sib.ch).
CC -----
DR EMBL; X75326; CAAS3075.1; -.
DR PIR; S43833; S43833.
DR HSSD; Q59931; 1E0H.
DR MaizeDB; 78926; -.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; Aldehdh; 1.
DR TIGRFAMS; TIGR01804; BADH; 1.
DR TIGRFAMS; TIGR01722; MMSDH; 1.
DR TIGRFAMS; TIGR01780; SSADH; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
DR KX NADP; Oxidoreductase.
FT NP BIND 247 251 NAD (ADP part) (By similarity).
FT ACT_SITE 266 266 By similarity.
FT ACT_SITE 300 300 By similarity.
SQ SEQUENCE 498 AA; 53146 MW; 7AF1C0DACA8EE39 CRC64;
Query Match 48.7%; Score 1164.5; DB 1; Length 498;
Best Local Similarity 49.0%; Pred. NO.1.4e-61;
Matches 234; Conservative 88; Mismatches 14; Indels 13; Gaps 5;
QY 5 YONYNGEWKSVN--QIEILSPIDDSSLGFWPMATREEDVHANKAGREALPAAALTVY 62
Db :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 18 YRYVADGEWMTSASGSVAIVNPTRTKQYRVQCTGEVKNKAMDAAVQAOKAWAPPLW 77
Db YRYVADGEWMTSASGSVAIVNPTRTKQYRVQCTGEVKNKAMDAAVQAOKAWAPPLW 77
QY 63 ERAQYLAKAADIIEBDKEELAVTLAKEISKAYNASVTEVVTADLIKYAEEGIRLSTSA 122
Db ERAQYLAKAADIIEBDKEELAVTLAKEISKAYNASVTEVVTADLIKYAEEGIRLSTSA 122
QY 78 KRADVLRKAAILKSHKAPIAECEVKEIAKPAKDAVSEVRSGDLVSYTEEGVRIIGE- 136
Db KRADVLRKAAILKSHKAPIAECEVKEIAKPAKDAVSEVRSGDLVSYTEEGVRIIGE- 136
QY 123 DEGGKM---DASTG---HKLAIVIRQPVVIATAPNYVPNNLSGSKTAPLLIGANYMF 176
Db ||::||::||::||::||::||::||::||::||::||::||::||::||::||::
QY 137 ---GLVVDSDSPFGEERNKYCLSSKTIPGVVLAIPEFPYPANLGSKITGPALLIGNALVL 193
Db ---GLVVDSDSPFGEERNKYCLSSKTIPGVVLAIPEFPYPANLGSKITGPALLIGNALVL 193
QY 177 KPPTGVSVCGLVLAFAFAEAGLPAGVFNTITGRGSELIDYVEHEBENFIIFTSSTVGQ 236
Db KPPTGVSVCGLVLAFAFAEAGLPAGVFNTITGRGSELIDYVEHEBENFIIFTSSTVGQ 236
QY 194 KPPTGVAVALAHMWHCFILAGPKGLISCVTGKSEIDPLVMHPGVNCISFTGG-DTGI 252
Db KPPTGVAVALAHMWHCFILAGPKGLISCVTGKSEIDPLVMHPGVNCISFTGG-DTGI 252
QY 237 RIQLKLAKRRPIMLEIGKGDAQIVLADADLDNAKOIVAGAAYDYSGQRCTAIRLVLYBEV 296
Db RIQLKLAKRRPIMLEIGKGDAQIVLADADLDNAKOIVAGAAYDYSGQRCTAIRLVLYBEV 296
QY 253 AI SKKAGNVPLQMEIGGDACVIEDADLDIVSNATVYGFSYSQORTAKAVVLTIMESI 312
Db AI SKKAGNVPLQMEIGGDACVIEDADLDIVSNATVYGFSYSQORTAKAVVLTIMESI 312
QY 297 ADELAETKSBNVAKLSVGDPPFNATVVVIDNSDFESLVVVAROKGAELNEFKRDG 356
Db ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
QY 313 ADAVAVQAKNAKIATKLKVGPREDSDITIVVLESSANFLIEGLVMDKKEGALFCOEYREG 372
Db ADAVAVQAKNAKIATKLKVGPREDSDITIVVLESSANFLIEGLVMDKKEGALFCOEYREG 372
QY 357 RLTPFGLEHDVTLDMKLAMEEPFGILDIRVKDKAEBAVAIALNKSDPGLQGSVTRDFQK 416

```

```

Db      373 NLIWPLLDLHDPDRRIAMBEPPFGLVPIRINSVYEEGIHHCNASHFGLOGCFTFDINK 4322
Qy      417 AFDIANKEVGTVHINNKTGRGDNPFPPLGLKSGAGVGGIYRYSIBAMTNVKSIVDM 474
Db      433 AILISDAMEITGVQINSAPRBPDPHFSFGGLDSDIGSGGITSINIMMTKRVSTVINTL 490

RESULT 22
GAPN_PEA ID GAPN_PEA STANDARD; PRT; 496 AA.
AC P81406;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9)
DE (Non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase)
DE (Glyceraldehyde-3-phosphate dehydrogenase [NADP+]) (Triosephosphate
DE dehydrogenase).
DN Name=GAPN;
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OC NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. Rosakrone; TISSUE=Shoot;
RX MEDLINE=94180387; PubMed=7545914;
RA Habenicht A., Hellman U., Cerff R.;
RT "Non-phosphorylating GAPDH of higher plants is a member of the
RT aldehyde dehydrogenase superfamily with no sequence homology to
RT phosphorylating GAPDH."
RL J. Mol. Biol. 237:165-171(1994).
CC -!- FUNCTION: Important as a means of generating NADPH for
CC biosynthetic reactions.
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + NADP(+) + H(2O)
CC = 3-phospho-D-glycerate + NADPH.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75327; CAA53076.1; -.
CC DR HSSP; Q59931; IEUH.
CC DR InterPro; IPR002086; Aldehyde_dehydr.
CC DR Pfam; PF001771; Aldehyd; 1.
CC DR TIGRFAMs; TIGR01804; BADH; 1.
CC DR TIGRFAMs; TIGR01722; MMSDH; 1.
CC DR TIGRFAMs; TIGR01780; SSADH; 1.
CC DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
CC KM Direct protein sequencing; NADP; Oxidoreductase.
CC FT NP BIND 245 249 NAD (ADP part) (By similarity).
CC FT ACT SITE 264 264 By similarity.
CC FT ACT SITE 298 298 By similarity.
CC FT ACT SITE 496 AA; 53004 MW; 52PADCGF77C613FA CR664;
CC
CC Query Match 48.5%; Score 1161.5; DB 1; Length 496;
CC Best Local Similarity 49.5%; Pred. No. 2.1e-61;
CC Matches 236; Conservative 85; Mismatches 145; Indels 11; Gaps 5
Qy      5 YONYNNGEMKSSVN--QIEIISPIDSSIGFVPAMTRREVDHDMKRGREALPMAALTGY 62
Db      16 YKYVADGEMKSTSGSKSVAILNPTTRKPYQYKQACSGEIVNKVMSAKAGSNMATTPLW 75
Qy      63 ERAQYIHRADAIIEBDEKEIATVLAKESIKAYNASVTEVTRADLIRVAEGIRLSTSA 122

```

```

Db      76 KBAELHKAALIKHEKAAIAECLVKEIAKPAKADATEVVRSGDLVSYCAEGRVRL---L 132
Qy      123 DGGGCK--DASTGH---KLAIVRRQVGVIVLAIAPNVPVNSGSLIAALLGANNVMMK 177
Db      133 GGGKFLVSSFPENERTKYCLTSKIPLVGVIALPPNVPVNLAVSKIAALLAGNSIVLK 192
Qy      178 PPTQGSVSGVLAKAPAEAGLPAGVFNITGTGSGSEIGDYIVEHEEYVNFNFNGSTFVQGR 237
Db      193 PPTQGAVALAHVHCHLAGFPKGLISCTYGKSEIGDPLTMHPGNCISFTGG-DTGIA 251
Qy      238 ICKLAGMRDIMELEGKDGAGIVLADADLNNAKQIVAGAYDYSQGRCTAIKRYLVVEEVA 297
Db      252 IKSXSGMIFLOMELGKDGACIVLEADLDLVANIIKGGFSYSQGRCTAVKVLVMEESA 311
Qy      298 DELAEKISENVAKLSTVGDPEDNATVTPVVDNSADTIESLVVDAROKAKELNEFRDGR 357
Db      312 DALVEKVKYKAVAKLSVGPEDSDITPVVSESSANFIEGLVMDAKKGAIFCOEYKREGN 371
Qy      358 LITPGLFDHVTLDMKLAMEEPPGPIPIIRVKDAEVAVALANKSPGLQSSVFTRDFOKA 417
Db      372 LITPGLLDHVRPDMRIAMEEPPGPVLPVIRINSVEEGIHHCNANSFGLQGVFTKDINKA 431
Qy      418 FDIANKLEVGTVHINNKTRGPDNPFPLGLKSGSAGVQGIIRYSIEAMTVKSYIVLDM 474
Db      432 IMISDAMESGTVOINSAPARGPDHPFPGQIKDSIGSGQGITNSINMWTIKVSTVINL 488

RESULT 23
Q629G0 PRELIMINARY; PRT; 499 AA.
ID 0629G0
AC 0629G0;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
GN Name=P0528B09.18; Synonyms=P0524F03.45;
OS Oryza sativa (Japanese cultivated group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RL Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004703; BAD09910.1; -.
DR EMBL; AF004702; BAD09896.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IP0002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd_1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
SQ SEQUENCE 499 AA; 53369 MW; 2477C6FB034DB34 CRC64;

Query Match 48.4%; Score 1158.5; DB 2; Length 499;
Best Local Similarity 49.0%; Pred. No. 3.2e-61;
Matches 224; Conservative 87; Mismatches 144; Indels 13; Gaps 5;
Qy 5 YONYVNGEWSVNV-QIIRILSPIDSSIGFVPAMTREEDHAMKAGREALPAMALITVY 62
Db 19 YRYVAGWEKVRVSSGSAIVINVTTLTQYVQACQOEENKMKMETAKYAQKAMATPIWM 78
Qy 63 ERAQVYHKAADIERDKEBIATVLAKEISKANNAVSTEVVTRADLIRYABEGIRLSTSA 122
Db 79 KBAELHKAALIKHEKTPIAECLVKEIKPAKADISEVRSGLVSYABEGVRILGR- 137
Qy 123 DGGGCK---DASTG---HKLAIVRRQVGVIVLAIAPNVPVNSGSLIAALLGANNVMMK 176
Db 138 ---GKLLVSDSPFGENRNTKYCLSSKVPVLGVIALPPNVPVNLAVSKIAALLAGNSIVLK 194
Qy 177 KPTQGSVSGVLAKAPAEAGLPAGVFNITGTGSGSEIGDYIVEHEEYVNFNFNGSTFVQGR 236

```

```

Db      195 KPTQGAVALAHVHCHLAGFPKGLINCVKSGSEIGDPLTMHPGVNCISFTGG-DTGI 253
Qy      237 RIGKLAGMRPIMELEGKDGAGIVLADADLNNAKQIVAGAYDYSQGRCTAIKRYLVVEEY 296
Db      254 AISKAGAVPLOMELGKDGACVLEADLDLVANIVYGGFSYSQGRCTAVKVLVMEESV 313
Qy      297 ADELAEKISENVAKLSTVGDPEDNATVTPVVDNSADTIESLVVDAROKAKELNEFRDGR 356
Db      314 ADIVEKVKAKIAKLTVPPEADSDITPVVSESSANFIEGLVMDAKKGAIFCOEYKREG 373
Qy      357 RLITPGLFDHVTLDMKLAMEEPPGPIPIIRVKDAEVAVALANKSPGLQSSVFTRDFOKA 416
Db      374 NLITPGLLDHVRPDMRIAMEEPPGPVLPVIRINSVEEGIHHCNANSFGLQGVFTKDINK 433
Qy      417 AFDIANKLEVGTVHINNKTRGPDNPFPLGLKSGSAGVQGIIRYSIEAMTVKSYIVLDM 474
Db      434 IMISDAMESGTVOINSAPARGPDHPFPGQIKDSIGSGQGITNSINMWTIKVSTVINL 491

RESULT 24
GAPN NICPL
ID GAPN NICPL STANDARD; PRT; 496 AA.
AC P93338;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9)
DE (Non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase)
DE (Glyceraldehyde-3-phosphate dehydrogenase [NADP+]) (triosephosphate
DE dehydrogenase).
GN Name=GAPN;
OS Nicotiana glumabiginifolia (leadwort-leaved tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=98036054; PubMed=9370287; DOI=10.1016/S0378-1119(97)00320-X;
RX Habenicht A., Quesada A., Cerif R.;
RA "Sequence of the non-phosphorylating glyceraldehyde-3-phosphate
RT dehydrogenase from Nicotiana glumabiginifolia and phylogenetic origin
RT of the gene family."
RL Gene 198:237-243(1997).
CC -!- FUNCTION: Important as a means of generating NADPH for
CC biosynthetic reactions.
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + NADP(+) + H(2)O
CC = 3-phospho-D-glycerate + NADPH.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U87848; AAB47571.1; -.
DR HSSP; O59931; IEUH.
DR InterPro; IPR002086; Aldehyd_dehydr.
DR Pfam; PF00171; Aldehyd_1.
DR TIGRFAMs; TIGR01804; BADH; 1.
DR TIGRFAMs; TIGR01722; MMSDH; 1.
DR TIGRFAMs; TIGR01780; SSADH; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
DR NADP; Oxidoreductase.
FT NP BIND 245 249 NAD (ADP part) (By similarity).
FT ACT SITE 264 264 By similarity.
FT ACT_SITE 298 298 By similarity.

```

SEQ	SEQUENCE	496 AA;	53143 MW;	6f350f168bd7a108 CRC64;
Query Match		48.3%;	Score 1155.5;	DB 1; Length 496;
Best Local Similarity		49.5%;	Pred. No. 4,8e-61;	
Matches 236;	Conservative	86;	Indels 144;	Gaps 11; Indels 11; Gaps 5;
OY	YONYNGEMKSSVN--QIEILSPIDSSLGFPNATREEDVHAMAGAEALPANAALTVY	62		
DB	16 FRYSEEGEMKSSASGSAVAIINPTTRKQYKQACTOEIVNMTQVATAOKSMVAKTPIW	75		
OY	63 BRAQVLIHRAADIIEEDKEIEATVLAKETISKVNAVSETEVPRADIIRYAAEEGLSTSA	122		
DB	76 KRAELIHRRAALIKKHKAPRIAECLYKEIAKPKADVTEVMSGDLVSTTAEGVRI---	132		
OY	123 DEGGKM--DASTGH--KLAVIRROPGVIGVLAIAEYNYPVNLGSGKIPALIGNVNFK	177		
DB	133 GEGKFLVDSFPGNERTKYCLTSKIPLVGILAIPEFNYPVNLAVSKIPALIAAGSLVLK	192		
OY	178 PPTQSSVSGVLAKAFPAEAGLPAGVFNITIGKSGEIGDIYHEHEVNFINTGTPVGQR	237		
DB	193 PPTQAVAVCLHMVHCFHLIAGFPKGLISCVTGKSGEIGDLTGMPEVHISFTGQ--DTGVA	251		
OY	238 IGLKAGMPIMLEGGKXAGIVLADADI.DNNAKO.VAAGAYDSGGRCAIRKRVYVESVA	297		
DB	252 ISKAGMPLQWELGGKXACIVLEADADIDLAAGSTVKGFSYGGRCVAVKVLVMSVSA	311		
OY	298 DELAEKISENVAKLSVDPFDNAVTVPIYIDNSADFISLIVDARQKAKELNEFKRDR	357		
DB	312 DALVKKVNAKAKALTVGPREDDCDITTPVSSSNFIEGLVMDAKQKATTCQYKREGN	371		
OY	358 LITPGLFPHVTLDMKLAAEEPPGPIILPIIRVKAEEAVALANKSDFGIQQSVFTEDFOKA	417		
DB	372 LIMPILLNVSPDMRIABEPPGPVLPIYIRINSVEEIGHNCAISNFGIQQCVFTDINKA	431		
OY	418 FDIANKLEVGTHINNTKGRGPDNPPIGLKSGAGCVGQIRYSIAAMNVKSYLIDM	474		
DB	432 ILISPMETGIVQINSAPARGPDPHFPFGIKDMSGIGSOGITNSIMMTKVTYIVINTL	488		
RESULT 25				
O864Y9				
ID	O864Y9	PRELIMINARY;	PRT;	507 AA.
AC	O864Y9;			
DT	01-JUN-2002 (Tremblrel. 21, Created)			
DT	01-JUN-2002 (Tremblrel. 21, Last sequence update)			
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE	Glyceralddehyde-3-phosphate dehydrogenase.			
OC	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Euharicotidae; Oryzae; Oryza.			
OX	NCBI_Taxid:4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=12582559;			
RA	Arumgam Pillai M., Lihuang Z., Akiyama T.;			
RT	"Molecular cloning, characterization, expression and chromosomal			
RT	location of OSGADH, a submergence responsive gene in rice (Oryza			
RT	sativa L.).";			
RL	Theor. Appl. Genet. 105:34-42(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Pillai A., Akiyama T.;			
RL	Submitted (Mar-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL, AF357884; NAM00227.1; --.			
DR	HSSP; Q59931; LEUH.			
DR	Gramene; O864Y9; --.			
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	InterPro; IPR02086; Aldehyde_dehydrog.			
DR	Pfam; PF00171; Aldedh. 1.			
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.			
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.			

SEQ	SEQUENCE	507 AA;	54206 MM;	15B706090463FFD5_CRC64;
QY	Query Match	47.9%;	Score 1145.5;	DB 2; Length 507;
Db	Best Local Similarity	48.4%;	Pred. No. 2e-60;	
Matches 235;	Conservative 88;	Mismatches 142;	Indels 21;	Gaps 7;
QY	5 YQNVNGEKMSVN--QIELSPIDSSLGFPVPMTEEDVDHAKAREALPMAALTYV 62	:	:	:
Db	19 YRYVADDEMRVSAGSKVAIVNPTTRLTOYRVQACTOEYVKNMETKVKQAKAARPLW 78	:	:	:
QY	63 ERAOYLKKAADIIRDKBEIATVLAKEISRYAVASVEYRTADLIRYAABEGRILSTSA 122	:	:	:
Db	79 KRAELHKAAILKEHKTPIAECLVKEIAPKADAISEVRSGLDVSYTIEBGRILGE- 137	:	:	:
QY	123 DEGGKM--DASTG--HKLAVIRROFVGIVLAIAPNYPVNTSGSKTIAPALIGSNVVF 176	:	:	:
Db	138 ---GKLVSDFPENENKTKCLSKVPLGLVVLAIPTPNYVNLAVSKIGPALIAGNALVL 194	:	:	:
QY	177 KPTQSGSVGLVLAKAFAEAGLPAGVFNITITGRGSEIGDYIVHEEYVNFNFTGSTVQ 236	:	:	:
Db	195 KPPTQGAVALAHMWHCFHLAGFPKGLINCVTGKSEISGDFLTMHPGVNCISFTGG-DTGI 253	:	:	:
QY	237 RIGTLAAGRPIMLTELGKDAQIVLADADLINAQOIVAGAVDVSQGCRTIKRLVYEEV 296	:	:	:
Db	254 AISKRAQVNPQWELGSKDACVLEDDLDLVAAINIVKGGYSYGQCTVAKVLLMESV 313	:	:	:
QY	297 ADELAEKISENVAKLTVGDPDFDNATVPVIDDNSADPISLSLVADAROKGA---KE----- 348	:	:	:
Db	314 ADIYEVKAKATLACTVTPPEADSDITPVYTSSANITBGLVMDAKKKGATPCKEKATP 373	:	:	:
QY	349 LNEFRDQRLITPGLFDHVTLLDMKLAMEEPPGPIILPIIRVKDAEVAVALANKSDPGLQSS 408	:	:	:
Db	374 QOEYRRGNGILWPLFLLDHVPDPMRIAMEEPPGVLPIVIRINSVEGIIHGNASVFLQGC 433	:	:	:
QY	409 VFTDPOKAPDIANKLEVGIVHINNTKGRPDNFPFGLKSGAGVCGIRYSIEAMTNVK 468	:	:	:
Db	434 VFTDINKALMISPMETGVQVINSAPARPDHPFGQLKDSGISGIGTINSINMTKVK 493	:	:	:
QY	469 SIVLDM 474	:	:	:
Db	494 STVINL 499	:	:	:
RESULT 26				
Q84PC4	PRELIMINARY;	PRT;	508 AA.	
AC	Q84PC4;			
DT	01-JUN-2003 (Tremblrel. 24, Created)			
DT	01-JUN-2003 (Tremblrel. 24, Last sequence update)			
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE	NADH-dependent glyceralddehyde-3-phosphate dehydrogenase			
DE	(Fragment).			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehretidae; Oryzae; Oryza.			
OX	NCBI_TaxID=39947;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=22584437; PubMed=12684538; DOI=10.1073/pnas.0737574100;			
RA	Cooper B., Clarke J., Budworth P., Kreps J., Hutchison D., Park S.,			
RA	Guttil S., Dunn M., Ludwig P., Billeo C., Goff S.A., Glazebrook J.;			
RT	"A network of rice genes associated with stress response and seed			
RT	development."			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:4945-4950(2003).			
DR	EMBL; AY224439; AA072558.1; -.			
DR	HSSP; Q59931; IEUH.			
DR	Graeme; Q84PC4; -.			
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	InterPro; IPR002086; Aldenhyd_dehydrog.			
DR	Pfam; PF00171; Aldedh; 1.			
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.			

DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 508 AA; 54082 MW; FF027ACBF2743F58 CRC64;
Query Match 47.6%; Score 1138.5; DB 2; Length 508;
Best Local Similarity 48.9%; Pred. No. 5.1e-60;
Matches 231; Conservative 85; Mismatches 143; Indels 13; Gaps 5;
OY 11 GEMKSVN--QIEILSPIDSSIGFVPAMTRESVDHAKGRBALPAMALTYERAQYL 68
DB 34 GMRVASAGSKVAIVPTKRLTQRYOACTQGEVNVKMTAKYAKQAMARTPLMKRAELL 93
OY 69 HKRAADIERDKEIATVLAKESIKAVNASVTEVTPADLIRYAABEGIRLSTADEGGM 128
DB 94 HKAAALIKHKTPRIAECLVKEIAKPAKDAISEVVRSGDLVSTAEGBVILDS---GKL 149
OY 129 ---DASTG---HKLAIVRRQPGVIVLAPVNYPVNLSSKTAIPALIGNVMEKPTQG 182
DB 150 LVSDSPFGNRRKYYCISSTKVPILGVLAIPFPVNVLAWSKIGPALIAGNALVLKPTQG 209
OY 183 SVSGVILAKAFABAGLPAGVFNTITGRGSEIGDYIYEHEVNFINTGSTPVQORIGKLA 242
DB 210 AVAALHMHVCHFLAGEPKKLINCVTGKSEIGDFLTMHGVNCISFTGG-DTGIAISKKA 268
OY 243 GMRPIMLEIGKDGAGIVLADADLDNAKQIVAGAVDYSQORCTAIRVLVBEVADAE 302
DB 269 GNVPLQMEIGKDGACVLEDDADLDVAANIYKGGFSYGORCTAVKVLIMESVADIYE 328
OY 303 KISENVAKISVGDPPDNATVTPVIDNSADFTESLIVDARQKAEKELNEFKDGLLTG 362
DB 329 KYAKKLAKLTVPPEADSDITPVVTESSANFLIEGLVMDAKEKATFCQBYRRGNLWPL 388
OY 363 LFDHVLDMKLAMEEPPGILPLIRYKDAEEAVAIANKDFGLQSSVFTFRDOKAPDIN 422
DB 389 LLDHVRPDMRIAMEEPPGVLPIVIRINSVEBGHHGNANPFLQCGVFTKDKINKAIMISD 448
OY 423 KLEVGTVHINNTGRGPDNFPPLGLKSGAGVQIGRYSIEAMTNVKSIVLDM 474
DB 449 AMETGVQINSAPARGPDHFPQGLKDSIGSGIGTNSIMMTKVSVTINL 500
RESULT 27
O6RCS1 PRELIMINARY; PRT; 496 AA.
AC O6RCS1;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Non-phosphorylating GAPDH.
GN Name=ALDH11A5;
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
OX NCBI_TaxId=3218;
RN [1]
RP SEQUENCE FROM N.A.
RA Wood A.J., Reeki R., Frank W.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY504666; AAS78753.1; -;
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd. 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
SQ SEQUENCE 496 AA; 53347 MW; AEB641DC6CE3A8A6 CRC64;
Query Match 46.8%; Score 1120.5; DB 2; Length 496;
Best Local Similarity 47.9%; Pred. No. 6e-59;
Matches 229; Conservative 89; Mismatches 147; Indels 13; Gaps 5;
OY 5 YQNVNNGEKK--SSVQVQIILSPIDSSIGFVPAMTRESVDHAKGRBALPAMALTYE 62
DB 16 FRYADGEWKVSSGSKSVGITNPSTLKQYKVOACTQDEVNKAIVESAQAOKIWAQTPIM 75

OY 63 ERAQYLHKAADIERDKEIATVLAKESIKAVNASVTEVTPADLIRYAABEGIRLSTSA 122
DB 76 KRAEALHFPALIKKKNKEIALVKEIAKHKALTEVNASGDLISABEGIRILAE- 134
OY 123 DEGGM---DASTG---HKLAIVRRQPGVIVLAPVNYPVNLSSKTAIPALIGNVME 176
DB 135 ---GKFLVSDSPFGNRRKYYCISSTKVPILGVLAIPFPVNVLAWSKIPALIAVIL 191
OY 177 KPTQGSVGLVLAFAFABAGLPAGVFNTITGRGSEIGDYIYEHEVNFINTGSTPVQ 236
DB 192 KPTQGAASVLMHVCMLMAGPPKGLVSAITGKSEIGDLTMHGVNCISFTGG-DTGI 250
OY 237 RIGKLARGPIMLEIGKDGAGIVLADADLDNAKQIVAGAVDYSQORCTAIRVLVBEV 296
DB 251 AISRKAQVPLQMEIGKDGCCVLEDDADLELANNVIGGYSQORCTAVKVICVMSV 310
OY 297 ADELAEKISENVAKISVGDPPDNATVTPVIDNSADFTESLIVDARQKAEKELNEFKDG 356
DB 311 ABEIVSKIVQKMTKLTVMPEDNCDITPVVQSANFLIQGLVEDAQAKAFHQEKREG 370
OY 357 RLTTGLDHTLDMKLAMEEPPGILPLIRYKDAEEAVAIANKSDPGLQSSVFTFRDOK 416
DB 371 NLIWPLLDVNTVPDMRIAMEEPPGVLPIVIRIKTYEBGHHGNANPFLQCGVFTKDKIN 430
OY 417 AFDIANKLEVGTVHINNTGRGPDNFPPLGLKSGAGVQIGRYSIEAMTNVKSIVLDM 474
DB 431 AILVSDAMESGTIQINAPARGPDHFPQGLKDSIGSGIGTNSIQMNTKTKSVTINL 488
RESULT 28
O6MT99 PRELIMINARY; PRT; 471 AA.
AC O6MT99;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.9).
GN Name=gapN; OrderedLocustNames=MSC 0509;
OS Mycoplasma mycoides (subsp. mycoides SC).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxId=44101;
RN [1]
RP SEQUENCE FROM N.A.
RA Westberg U., Petersen A., Holmberg A., Goessmann A., Lundberg J.,
RA Johansson K.-B., Petersson B., Uhlen M.;
RT "The genome sequence of Mycoplasma mycoides subsp. mycoides SC type
RT strain PGIT, the causative agent of contagious bovine pleuropneumonia
RT (CBPP).";
RL Genome Res. 14:221-227(2004).
DR EMBL; EX842643; CAE77137.1; -;
DR HSSP; P20000; IAG8.
DR GO; GO:0008866; F:glyceraldehyde-3-phosphate dehydrogenase (N. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldehyd. 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW complete proteome.
SQ SEQUENCE 471 AA; 51976 MW; 5933A820300770ED CRC64;
Query Match 45.1%; Score 1079; DB 2; Length 471;
Best Local Similarity 44.7%; Pred. No. 1.7e-56;
Matches 210; Conservative 108; Mismatches 150; Indels 2; Gaps 1;
OY 4 EYQNVNNGEKKSVNQIILSPIDSSIGFVPAMTRESVDHAKGRBALPAMALTYE 63
DB 3 KFKALLDGLPDNNRIEILINVPDSVAGVSLTKQDINAFIAKSSQKAMENTDLEK 62
OY 64 RAQYLHKAADIERDKEIATVLAKESIKAVNASVTEVTPADLIRYAABEGIRLSTAD 123

```
DB 63 RISILDKKQILIDONKEELAQIIMSEFAKYKDCLEFVIRSVYIQTFFVEKNLTLII 122
QY 124 EGGRMDASTGHKLAVIRROPVGIVALIAPYVPVNLSGSKLAPALIGNVVMKPPTOGS 183
DB 123 DGAKYGA--NKIGTFMRVAKGVGVAISPENYINLAVSKIPLCLVGNITVFKPATQGS 180
QY 184 VSGVLVAKAFAGLPAAGVNTITGSGSEIGDIYVEHEVNFNFPGSTPVGQRIGLAG 243
DB 181 LIGAKGELAYQNLKPGIFNVVYTGREGREIGDIIITNKLADFLSFGSVEVGRLEIIS 240
QY 244 MRPIMLEGGKDGIVLADADLDNAKOIYAGAVDYSGORCTAIKRVLVVEEVADELAEK 303
DB 241 TKVAVLEGGKDAIYVLDLDLEKRYAKEIISGAFSISGQRCIAKRVITTDKADQLVPL 300
QY 304 ISENVAKLSVGDPEPDNAVTVTVIDNSADPIESLVVDAROKAKEINEFRDGRLLTPGL 363
DB 301 LKEKINKLTVGLPKDNCIDITPLIDOKTADPFVGLIDAKKKAKIIGDKQKNLITPPL 360
QY 364 FDHVTYLDMLKAMEPPFPPIPIIRVYDAEBAVAIANKSDPGLQSSVFTPDFOKAFDIAK 423
DB 361 VDHVTDMLAMEPPFPVPIIRITSVDQMIELANKSNFGLQASVYTKLDDALTYIAOK 420
QY 424 LEVGTVINNKTRGPDNPFPPLGLKSGAGVQIGIRYSIEAMTVKSIYVL 473
DB 421 LEVGTVINNKSGORGPDPVFPPLGVKDSGFVGQIYVDTLFLTRYKGIYIN 470
RESULT 29
Q6FLK7 PRELIMINARY; PRT; 472 AA.
AC Q6FLK7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)
DE NADP-dependent glyceroldehyde-3-phosphate dehydrogenase.
GN Ordered locus names=M1259;
OS Mesoplasma florum (Acholeplasma florum).
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Entomoplasmataceae; Mesoplasma.
NCBI_TaxID=2151;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H1 / ATCC 33453;
RA Birten B.W., Strange-Thomann N., Hafez N., Decaprio D., Fisher S.,
RA Butler J., Birkins T., Kodira C.D., Major J., Wang S., Nicol R.,
RA Nusbhum C.;
RL Submitted (JTN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL:AE017263; F:oxidoreductase activity; IEA.
DR GO:GO:0016491; P:oxidoreductase activity; IEA.
DR GO:GO:0008152; P:metabolism; IEA.
DR InterPro:IPR002086; Aldehyd_dehydrog.
DR Pfam:PF00171; Aldehyd. 1.
DR PROSITE:PS00070; ALDEHYDE DEHYDR_CYS; UNKNOWN_1.
DR PROSITE:PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Complete proteome.
SQ
SEQUENCE 472 AA; 51718 MW; 66EFA081AC4E8C04 CRC64;
Query Match 45.0%; Score 1077.5; DB 2; Length 472;
Best Local Similarity 44.8%; Pred. No. 2.1e-56;
Matches 211; Conservative 110; Mismatches 147; Indels 3; Gaps 2;
QY 3 KEIYQNVNGEKKSVNOIEILSPIDSSLGFPVPMTRREEDVDMKAGREALPMAALTVY 62
DB 2 EKKAIINNGEISNSGLIDMDPTDEVVAQVSAISQEIISDAFKAKAAQKQKMAIGIE 61
QY 63 ERQYVILHKADIERDKKEBIATYLAKEISKAYNASTVEVVRTDLIRYAEGRIST-8 121
DB 62 KRTEFLIRMDLLKKEEDLATTMHEIARAYKDCLEVRRTAEYIDLITSEYNDLQVLT 121
QY 122 ADEGKMDASTGHKLAVIRROPVGIVALIAPYVPVNLSGSKLAPALIGNVVMKPPPTO 181
DB 122 FDKNSK--GVTEIDIVAEYKRIAGGVGVSIPPIYPIINLAVSKLAPGLTNTVFEPKAT 179
```

```
QY 182 GSVSGVLVAKAFAGLPAAGVNTITGSGSEIGDIYVEHEVNFNFPGSTPVGQRIGKL 241
DB 180 GSVVGIKIGQLALEGIPAGVNLVYTGREGRIDGVITVNPPLIDFLSFTSVPGRLEIMBI 239
QY 242 AGMRPIMLEGGKDGIVLADADLDNAKOIYAGAVDYSGORCTAIKRVLVVEEVADELA 301
DB 240 SSSKDLVLEGGKDAIYVLDLDLEKRYAKEIISGAFSISGQRCIAKRVITTDNDKTT 299
QY 302 EKISENVAKLSVGDPEPDNAVTVTVIDNSADPIESLVVDAROKAKEINEFRDGRLLTP 361
DB 300 PLIKAEVAKLTVGLPENENIITPMDIKSADPFVDTLINDAVKGAFLVYGTRDKNLLOP 359
QY 362 GFDPHTYLDMLKAMEPPFPPIPIIRVYDAEBAVAIANKSDPGLQSSVFTPDFOKAFDIA 421
DB 360 TLIDNVYVMNAVMEPPFPVPIIRINIEBMSIVANKSNFGLQSIYSKVDLAVKYVA 419
QY 422 NKLVEGTVINNKTRGPDNPFPPLGLKSGAGVQIGIRYSIEAMTVKSIYVL 472
DB 420 EQLEGTVINNKTRQRPDVLPLGLVQDSGFVGQIKETILSTTRYKGIIT 470
RESULT 30
Q8EVT9 PRELIMINARY; PRT; 478 AA.
AC Q8EVT9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE NADP-dependent glyceroldehyde-3-phosphate dehydrogenase.
GN Ordered locus names=MYPB4710;
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22354719; PubMed=1246555; DOI=10.1093/nar/gk1667;
RX Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5283-5300(2002).
DR EMBL:AP004172; BAC44260.1; -.
DR HSSP; Q59931; 1EDH.
DR GO:GO:0016491; P:oxidoreductase activity; IEA.
DR GO:GO:0008152; P:metabolism; IEA.
DR InterPro:IPR002382; Aldehyd_dehydrog.
DR InterPro:IPR002332; PII_GlnB_UMP_S.
DR Pfam:PF00171; Aldehyd. 1.
DR PROSITE:PS00070; ALDEHYDE DEHYDR_CYS; UNKNOWN_1.
DR PROSITE:PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR PROSITE:PS00496; PII_GlnB_UMP; UNKNOWN_1.
KW Complete proteome.
SQ
SEQUENCE 478 AA; 52311 MW; ADDA6110273087C5 CRC64;
Query Match 43.8%; Score 1047.5; DB 2; Length 478;
Best Local Similarity 46.5%; Pred. No. 1.3e-54;
Matches 218; Conservative 90; Mismatches 152; Indels 9; Gaps 3;
QY 8 YNGEKKSVNOIEILSPIDSSLGFPVPMTRREEDVDMKAGREALPMAALTYVERAOY 67
DB 11 YNGKRVSEKTIPIVSTIDGSLIGVVALSKKIDIMFEAGHLFNLNRNLTPKWRIRK 70
QY 68 LHKADIERDKKEBIATYLAKEISKAYNASTVEVVRT----ADLIRYAEGRISTSD 123
DB 71 IKEPAEYFAEKEPFLTLMSYELGKSYNDALKEVERSYNIFETIEVYKEKEVNSVIGP 130
QY 124 EGGRMDASTGHKLAVIRROPVGIVALIAPYVPVNLSGSKLAPALIGNVVMKPPTOGS 183
DB 131 EVNKKIKGKTGYVYVY---PVGVLVLAIPFNPYINLGLAKIIPILLVGNITVFKPATQGS 186
QY 184 VSGVLVAKAFAGLPAAGVNTITGSGSEIGDIYVEHEVNFNFPGSTPVGQRIGKLAG 243
```

```
Db      : ||| :| |||| :||| ||||| :| :| |||| :| :|
187 IYSSOLAKYFDQANFAGVFNLTGKSGSEIGDIYLENKRIQATFTGSSDIEFKLASKLP 246
Qy      244 MBPIMLEIGCGKAGIYVLADADLDNNAKOIVAGAYDVSQGRCTAIRKRVLVVEADLEAK 303
Db      247 MBPIMLEIGCGKAAIYVNNADVELAKEIITKAPGVSQGRCTAIRKRVLTKEIADLLIH 306
Qy      304 ISENAKLSVGDPPDNATVPYIDNSADFIESLVADARQAKELNEFKRG-RLTFPG 362
Db      307 LIKEAQKLVGNPLANNPDITPLIDRAVDVFNIDLDALNHGANLVYGINIGYNIILRHT 366
Qy      363 LFDVATLDMKLAWEPPFGPIPIIRVKAEEAVAIANKSDPGLQSSVFTFRDPOKAFDIAN 422
Db      367 IIDNVSTSKLAMEEPPFGVPLPIIRVSTINEAIAIINASEYGLQSGVFTFKDLEBARTIAK 426
Qy      423 KLEVGTVHNNKTGSGPDNPFPLGLKSGAGVQGRYSIEAMTNYSIV 471
Db      427 YLDGTVINIRGSSRQPDIFPFIITKSGFGVQGRISDKAMTRVGIIV 475
```

RESULT 31

```
Q8LK61 PRELIMINARY; PRT; 496 AA.
ID      Q8LK61
AC      Q8LK61;
DT      01-OCT-2002 (TREMBlrel. 22, Created)
DT      01-OCT-2002 (TREMBlrel. 22, last sequence update)
DE      01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE      Nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (EC
1.2.1.9).
OS      Triticum aestivum (wheat).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC      Triticeae; Triticum.
OX      NCBI_TaxId=4565;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=leaf;
RX      MEDLINE=22275840; PubMed=12387887; DOI=10.1016/S0014-5793(02)03455-5;
RT      Bustos D.M., Iglesias A.A.;
RT      "Non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase is post-
translationally phosphorylated in heterotrophic cells of wheat
(Triticum aestivum)."
RL      FEBS Lett. 530:169-173(2002).
DR      EMBL; AF521191; AAM77679.1; -.
DR      HSSP; Q59931; IEUH.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR002086; Aldehyd_dehydrog.
DR      Pfam; PF00171; Aldehyd; 1.
DR      PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW      Oxidoreductase.
SQ      SEQUENCE 496 AA; 52872 MW; 20F9D07F1F840FB CRC64;
```

Query Match 42.9%; Score 1027.5; DB 2; Length 496;
Best Local Similarity 45.0%; Pred. No. 2, 2e-53;
Matches 215; Conservative 78; Mismatches 172; Indels 13; Gaps 4;

```
Qy      5 YQNVNNGEKKSSVN--QIILISPIDSSIGFVPAMTRREVDHAMKAGEBALPAMALTVY 62
Db      16 YKYVADGEWRASASGKTVAIINPTTRQTYRVOACTQOEBVNKMDAKVAKXKWATPTLM 75
Qy      63 EBAQVILHKAADIIEKDKEIATVLAKEISKAYNASVTEVTRADILIRYAABEGIRLSTSA 122
Db      76 KRAELIHKAAALIKHEKTPIAECLVKEIAKPAKDAVSEVVRSDLVSTYAEBSVRLIGE- 134
Qy      123 DEGGKMDASTG-----HKLAVIRQPGVIVLAIPYNYPVNLSSGIAPALIGNVNMF 176
Db      135 ---GKLXVNSPFGNERNKYCLSSKIPLGVALIIPFPYIPANLASKIGPALIAGNALVL 191
Qy      177 KPPTGVSGLVLAKEFAAGAPAGVFNITITGRGSEIGDIYVHEEVNFINFTGSPVQO 236
Db      192 KPPTGAAVALHMHVCHFLAGFPKGLISCVTGKSGSEIGDPLTMHPGVNCISFTG-DTGI 250
```

```
Qy      237 RIGKLAKGRPIMLEIGCGKAGIYVLADADLDNNAKOIVAGAYDVSQGRCTAIRKRVLVVEEV 296
Db      251 AISKAGAVVPLXMKLXGDACIVXEXADLDLVXANIVVGGRFSYSGQRCAVNAKVVILIMEXX 310
Qy      297 ADELAEKISENAKLSVGDPPDNATVPYIDNSADFIESLVADARQAKELNEFKRG 356
Db      311 ADVXVZKVNAKLAKIKVPPPEXXXDITVVTSSANFVXGGBXCKGRXNFLPVYQBGK 370
Qy      357 RLTFPGLDVHTLDMKLAWEPPFGPIPIIRVKAEEAVAIANKSDPGLQSSVFTFRDPOK 416
Db      371 XLIMPDLHDVRRPDKXIAMKEPPGFVPLVIRINSVEBGIHHCNNSFGLQCGVFTFRDINK 430
Qy      417 AFDIANKLEVGTVHNNKTGSGPDNPFPLGLKSGAGVQGRYSIEAMTNYSIVLDM 474
Db      431 AIMISDAMESGTVOQNSAPARGPDHFPFGKDKSGIGQGITNSINMTTKVSYVINL 488
```

RESULT 32

```
Q8LSJ9 PRELIMINARY; PRT; 496 AA.
ID      Q8LSJ9
AC      Q8LSJ9;
DT      01-OCT-2002 (TREMBlrel. 22, Created)
DT      01-OCT-2002 (TREMBlrel. 22, last sequence update)
DE      01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE      Nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (EC
1.2.1.9).
OS      Triticum aestivum (wheat).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC      Triticeae; Triticum.
OX      NCBI_TaxId=4565;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Endosperm;
RX      MEDLINE=22275840; PubMed=12387887; DOI=10.1016/S0014-5793(02)03455-5;
RT      Bustos D.M., Iglesias A.A.;
RT      "Non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase is post-
translationally phosphorylated in heterotrophic cells of wheat
(Triticum aestivum)."
RL      FEBS Lett. 530:169-173(2002).
DR      EMBL; AF521190; AAM77678.1; -.
DR      HSSP; Q59931; IEUH.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR002086; Aldehyd_dehydrog.
DR      Pfam; PF00171; Aldehyd; 1.
KW      Oxidoreductase.
SQ      SEQUENCE 496 AA; 52869 MW; 8C1B1F15FF33881B CRC64;
```

Query Match 42.9%; Score 1026.5; DB 2; Length 496;
Best Local Similarity 45.0%; Pred. No. 2, 5e-53;
Matches 215; Conservative 78; Mismatches 172; Indels 13; Gaps 4;

```
Qy      5 YQNVNNGEKKSSVN--QIILISPIDSSIGFVPAMTRREVDHAMKAGEBALPAMALTVY 62
Db      16 YKYVADGEWRASASGKTVAIINPTTRQTYRVOACTQOEBVNKMDAAKVAKXKWARTPLM 75
Qy      63 EBAQVILHKAADIIEKDKEIATVLAKEISKAYNASVTEVTRADILIRYAABEGIRLSTSA 122
Db      76 KRAELIHKAAALIKHEKTPIAECLVKEIAKPAKDAVSEVVRSDLVSTYAEBSVRLIGE- 134
Qy      123 DEGGKMDASTG-----HKLAVIRQPGVIVLAIPYNYPVNLSSGIAPALIGNVNMF 176
Db      135 ---GKLXVNSPFGNERNKYCLSSKIPLGVALIIPFPYIPANLASKIGPALIAGNALVL 191
Qy      177 KPPTGVSGLVLAKEFAAGAPAGVFNITITGRGSEIGDIYVHEEVNFINFTGSPVQO 236
Db      192 KPPTGAAVALHMHVCHFLAGFPKGLISCVTGKSGSEIGDPLTMHPGVNCISFTG-DTGI 250
Qy      237 RIGKLAKGRPIMLEIGCGKAGIYVLADADLDNNAKOIVAGAYDVSQGRCTAIRKRVLVVEEV 296
Db      251 AISKAGAVVPLXMKLXGDACIVXEXADLDLVXANIVVGGRFSYSGQRCAVNAKVVILIMEXX 310
```

```
OY      2 57 ADELAEKISENNAKUSVDPFPDNATYTPVIDNSADPFISLVVDAPQKAKELNEFKRDG 356
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      3 11 ADXVVZKVNAKLAKLKVGPPEXXXDITXTVTTESSANFIKSGGBXCXGERNFPLPVQEKG 370
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      3 57 RLUTGFLPHVTLDMKLANEPPPGPLPIPIIRKYDAEAVALANKSPFGLOSSFTFDFOK 416
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      3 71 XLIWPLLLDHVPRDDXXIAXKEPFGPALVIARINSVEEGJHCNASNFGLOGCVFTRDINK 430
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      4 17 AFDIANKLEVGIVHINNTGRGPDNPPLGLKGSGAGVOGIRSIEMAMTVYSIVLDM 474
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      4 31 AIMIDAMESGVJOINSAPRGDHPRFGGLSDSGSOGITNISIMMKTVKSVTINV 488
```

RESULT	33
07NBXS	
ID	07NBXS
AC	07NBXS;
DT	01-MAR-2004 (TREMBLrel. 26, Created)
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	PfA.
GN	ORFNames=MGA_0860;
OS	Mycoplasma gallisepticum.
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NCBI	NCBI_Taxid=2096;
RA	[1]
RP	SEQUENCE FROM N.A.
RC	STRAINER;
FX	MEDLINE=22830409; PubMed=12949158; DOI=10.1099/mic.0.26427-0;
RA	Padanil L., Garton T.S., Kuchish G., Marham P.F., Browning G.F.,
RA	Nguyen D.K., Swartzell S., Madan A., Mahairas G., Geary S.J.;
RT	"The complete genome sequence of the avian pathogen Mycoplasma
RT	gallisepticum strain R(1ow)";
RL	Microbiology 149:2307-2316(2003).
DR	EMBL; AE016967; AAP56485.1; -.
DR	HSSP; Q59931; 1E0H.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0008152; P:metabolism; IEA.
DR	InterPro; IPR002086; Aldehyde_dehydrog.
DR	Pfam; PF00171; Aldehyd. 1.
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
FW	Complete proteome.
SEQ	SEQUENCE 525 AA; 58881 MW; 9DA7CF18215BA981 CRC64;

Query Match	39.4%;	Score 944;	DB 2;	Length 525;
Best Local Similarity	45.0%;	Pred. No. 2.3e-48;		
Matches 210; Conservative	76;	Mismatches 173;	Indels 8;	Gaps 5;

```

Qy 9 VNGEKKSSVNOEILISPIDDSSLGVPAMTREVVDHAKMAGRRLALPAM-ALTVERACQ 67
Db 60 INNOVGSSKLEISPIDHEISGSFYGMATAEIEDDAYEADIAFKSSWSAGYEYKAKI 119
Qy 68 LHKADIIERDEEBEATVLAKEISKAYNASVTEVVPADLR--YAAEKGILSTSADE 124
Db 120 L-KPAQLLDHKEEFANILLTDNIATAKYNSSLAEVTRSIGYIHDTISVEEMIHPLWDE 178
Qy 125 GGRMASTGHKLAVIRROPVGLVLAIAFNTYPNLSSGSKIAPLIGANVYMFEPPIQSV 184
Db 179 --KIH-RVKEKXGKVFREPLGVALNISPNNYPLNTPLSKOITPYLIGNTVVYVAATOAL 235
Qy 185 SGLVLAKEFAEAGLPGAGVNTITGRGSEIGDGYVEHEEVLNIFPTSGSTPVGRIKGLAM 244
Db 236 IGRILAHLEFKQAGPBGVGVQCVVAGGREIGDKLNTMKHKISIFSTGSTPVGSLKQSAV 295
Qy 245 RIMLELGKQKAGIYIADADLNMAQOYIAGVADVSGGQCTAIKRLVYVEVDLAELAKI 304
Db 296 GNISLELGKQDALVLSIDYDQKOTTEIYKGAISYSGORCTAIKRVFVNKDEDEVOQL 355
Qy 305 SENVAKLSVGDPEFDNATVPVVIDNNSADFTESLIVAROKAKELNEFKDGRLLTPGLF 364
Db 356 KEVEKELHNGPNPNNSLVPVDOHNAANTIKELFDDAVTKCALVILGDEPKUCVEAILL 415

```

QY 365 DHTLDMKAMEEPPGPIPIPIIRVDAAEVAIVANKSPFGIOSSVFRDQKAFDIAKTL 444
Db 416 DHVTTTMRVAMEEPPGPIPIPIIRVNEAMEVKNHNSOYSGIQAISIFPSDQKDELANOL 475
QY 425 EVGTVAHNKTKGRGPNPEPFLGKSGAGVOGIRYSIEAMNVASIV 471
Db 476 EAATITNNRRSSRGPDPEPFPMGVKOSGVOGIGIRITLTISVRYRGFV 522

RESULT 34			
Q9PQC9	PRELIMINARY;	PRT;	475 AA.
AC Q9PQC9;			
DT 01-OCT-2000 (TrEMBLrel. 15, Created)			
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE MAP-dependent glyceralddehyde 3-phosphate dehydrogenase.			
GN Name=gapN; OrderedLocustNames=U0362;			
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).			
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.			
OX NCBI_TaxID=134821;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=ATCC 700970;			
RA MEDLINE=20500219; PubMed=11048724; DOI=10.1038/35037619;			
RX Glass J.I., Leftkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,			
RA Cassell G.H.;			
RT "The complete sequence of the mucosal pathogen Ureaplasma			
RT urealyticum."			
RL Nature 407:757-762(2000).			
RL EMBL; AE002133; AAF30771.1; -.			
DR HSSP; Q59931; 1EUH.			
DR GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR GO; GO:0008152; P:metabolism; IEA.			
DR InterPro; IPR002086; Aldehyd_dehydrog.			
DR Pfam; PF00171; Aldehyd_1			
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.			
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.			
DR Complete proteome.			
FW SEQUENCE 475 AA; DDB0D06576603752 CRC64;			

Query Match	36.3%;	Score 868.5;	DB 2;	Length 475;
Best Local Similarity	39.7%;	Pred. No. 6.8e-44;		
Matches 190; Conservative	95;	Mismatches 177;	Indels 17;	Gaps 5

[illegible]


```

DT 01-MAR-2002 (T:EMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (T:EMBLrel. 26, Last annotation update)
DE Aldehyde dehydrogenase.
GN OrderedLocustNames=PAE2480;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
NCBI_TaxId=13773;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=1M2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Flitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum."
RL EMBL; AEO09879; AL64225.1; -.
DR HSSP; Q59931; IEUH.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR Pfam; PF00171; Aldehdh; 1.
KW Complete proteome.
SQ SEQUENCE 478 AA; 52861 MW; 9FC9409116223E7 CRC64;

Query Match 32.8%; Score 786; DB 2; Length 478;
Best Local Similarity 37.1%; Pred. NO. 5.9e-39;
Matches 176; Conservative 96; Mismatches 192; Indels 10; Gaps 4;

QY 6 QNTVNGEWSKSNQVQIETLSPIDSSICGFVPMTRREEDHMKAGREALPMALTYERA 65
DB 5 KNINFEIIVERERRPHSHSIDSVIAETVSKEEDKAAIDAYDLKMSQULPAIKRA 64
QY 66 QYLHAKADIIERDKKEIATYLAKEISKANVASYTEVTRADLIRYAEBCIRLSTADEG 125
DB 65 EYUUKLEYIKSRBEELINILMVEGGIYKKWGEVFTRLVRNAAE-----LARHYG 118
QY 126 GK-MDASTGKHLAVIRROPYIGYLAIPVYVYNLGGSKTAPLIGNVMPFPPOGVS 184
DB 119 GKTLQSPAEQVSVNMFRRPGVGVITPMVYPLISMKKTAHMLATGNTIIVYRPASTPV 178
QY 185 SGLVLAKAFPAEAGLPAGVFTITGRGSEIGDYVHEHEVFNIPFGSTPVGQRIKLAG- 243
DB 179 TGRKIAMTHQAFPPKQVFLVVGPGSTVGEDEVVANKKSHVFTGESATGRILASRAGS 238
QY 244 -MRPIMLELGKQDAGIVLADADVADNNAKQIVAGAYDYSQRCYAIKRVLVVEEVADELAE 302
DB 239 ALKVTVLEGGSDPLIITLADADINLAVRIVFGAFPHQGICTSSKRILVHERVYDVFK 298
QY 303 KISEENAKLSVGPFPDNAVTT-VIDDNGADFIESLVVDAROKAKELNEFKDGRLLTP 361
DB 299 KPEERVSQLEKVGDPBRRTVQGLISPRQADEMERFYQDVSVSGRLTGGKRGAYFWF 358
QY 362 GLFDHTLDMKLAMEEPFGPLPIIRYKDAEVAVALANKSDPGLQSVFTRDQKAFDIA 421
DB 359 TVVYDVDRNRIRKEEYFGPIRPVVKNDDEAVANDTEYGSAAVVTNTNRAFKIA 418
QY 422 NKELEGVTHINNTKGRPPNPFPLGLKSGAGVQGIARYSIEMNTNYSIVLDMK 475
DB 419 EATESGVVHINDVTMLAESHVPPGGIKASFGREGGWSVHEHTYDRMLITILR 472

RESULT 38
DHAI_BACSU STANDARD; PRT; 488 AA.
AC P42236;
DT 01-NOV-1995 (Rel. 32, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DE Probable aldehyde dehydrogenase ycdB (EC 1.2.1.3).
GN Name=ycdB; OrderedLocustNames=BST02470;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxId=1423;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95219079; PubMed=7704254;
RA Ogawa K.-I., Akagawa N., Nakamura K., Yamane K.;
RT "Determination of a 2154 bp nucleotide sequence around the 24 degrees
RT region of the Bacillus subtilis chromosome."
RL Microbiology 141:269-273(1995).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borisier R., Boursier L., Brans A., Brun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Guim S.Y., Glaeser P., Goffeau A., Golligly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Hatach J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaeer-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Konigstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.C.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Roche E., Roche B.,
RA Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Teipstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarelli A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Wellensgger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasunoko K., Yata K., Yoshida K., Yoshikawa H.F., Zumestein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -1- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC -1- PATHWAY: Ethanol utilization; second step.
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D30808; BAA06468.1; -.
CC EMBL; Z99105; CAB12041.1; -.
CC PIR; G69752; G69752.
CC
CC HSSP; P51977; 1BX8.
CC
CC Subtilisin, Bg1159, ycdB.
CC InterPro; IPR002086; Aldehyde_dehydr.
CC Pfam; PF00171; Aldehdh; 1.
CC TIGRPFAMs; TIGR01804; BADH; 1.
CC TIGRPFAMs; TIGR01722; MMSDH; 1.
CC TIGRPFAMs; TIGR01780; SSADH; 1.
CC
CC PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
CC Complete proteome; Hypothetical protein; NAD; Oxidoreductase.
KW NP BIND 233 NAD (ADP part) (By similarity).
FT ACT_SITE 255 255
FT ACT_SITE 289 289
FT ACT_SITE 289 289
FT By similarity.

```

```

SQ SEQUENCE 488 AA; 52414 MW; 9C37B32EF51D59B2 CRC64;
Query Match 31.9%; Score 763; DB 1; Length 488;
Best Local Similarity 36.9%; Pred. No. 1.4e-37;
Matches 170; Conservative 94; Mismatches 179; Indels 18; Gaps 6;

QY 5 YQNVNGEM--KSSVNOIILSPID--DSSLGFPVPMNTREVDHAKAGREALPMAALTV 61
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 10 YLNFINGEMVWKSQSGMVKVENDPADVNDIVGQNSTADVERAYVAANEAKTAMKRLNG 69
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 YERAOYLHKAADIIERDKSEIATVLAKESKAYNASVTEVPTADILIRAAEEGRLSTS 121
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 70 AARGOYLTKYADIMEERLEIAPACATREWGKTLPEAKGETARGIATILRYAAGEGKEXT-- 127
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 ADEGGMDASTGKHLAVIRROPVGIYLAIPYVYNLSGSKIPALIGANVMPFPPTQ 181
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 128 ---GDVTPSTDKDALMFTTRVPLGVGVISPMNFPALIPYKAPALVIGNVYIKPATE 184
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 182 GSVGLVIAKAPFAEAGLPAGVFNTITGRGSEIGDYIYEHEVNFINTGSTPVGRIKTL 241
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 185 TAVTCAKIIACFEAEGLPAGVINLVTPGQSVVGQGLAEHDGVAAVFTGSGNQGIKILGA 244
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 242 AGCR--PIMLEGGKAGIYIADADIDNAKOIVAGAYVYSGQRCYAIKRVLYVEVADE 299
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 245 ALARGAQYOLEMGKKNPVIVADADLEAAAEAVITGAFPSGTQKCTATSRVIVQSGIYER 304
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 300 LAEKISENVAKLSVGDPF--DNATVTPVIDDNSADFIETSLVVDAROKAKEL-----NE 351
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 305 FPEKLLQRTKDDITIGSLKEDVWMPYIAKKNQDNCLESTIEKKQKQGAULLIGEGLENG 364
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 352 FPRDGLLTPGLFDHVTLLDKLAMEBPFPIIPIIRVKDAEEAVALANKSDPGLQSSVPT 411
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 365 KYONGYVQPAIFPDNTSEMTIAQEEIFGPIVLTIVDSIEBALNIAIVDKGLASISFT 424
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 412 RDOFQAFDIANKLAEVCTVYHINNTKGRGPNFPFLGLKSGSA 452
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 425 ENIGRLSFIIDEIDAGLVIRINAESAGVELQAPPGKMKQSSS 465
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 39
YELL_METUA STANDARD; PRT; 463 AA.
ID YELL_METUA
AC 058806;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical aldehyde-dehydrogenase like protein MJ1411 (EC 1.2.1.-).
GN OrderedlocusNames=MJ1411;
OS Methanococcus jannaschii;
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxId=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geochagen N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Utecherback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RT J. Bacteriol. 187:2128-2138 (1995).
RT -i- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DB EMBL; U67561; AA899418.1; -.
DB PIR; B64476; B64476.
DB HSSP; Q59931; 1E0H.
DB TIGR; MJ1411; -.
DB Interpro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00172; Aldehyd. 1.
DR TIGRfam; TIGR01804; BADH; 1.
DR TIGRfam; TIGR01780; SSADH; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; FALSE NEG.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; FALSE NEG.
KW Complete proteome; Hypothetical protein; Oxidoreductase.
FT ACT_SITE 240 By similarity.
FT ACT_SITE 274 By similarity.
SQ SEQUENCE 463 AA; 51135 MW; 46F26568595B6A5 CRC64;

Query Match 31.7%; Score 759.5; DB 1; Length 463;
Best Local Similarity 37.1%; Pred. No. 2.2e-37;
Matches 175; Conservative 95; Mismatches 183; Indels 19; Gaps 8;

QY 8 YVNGEWSKSVNO--LEIISPIDSSLGFPVPMNTREVDHAKAGREALPMAALTVYERA 65
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 2 FIDGK---INREDMDVINPYSLVYIKKI PALSRSEAEKALDIAKYKEMKMLPTTKRY 58
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 66 QYLHKAADIIERDKSEIATVLAKESKAYNASVTEVPTADILIRAAEEGRLSTSADEG 125
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 59 NILNIAKQIKKEKELKILAIIDAGKPIKQARVEVERISGTFKLA--FYVEKHREIV 115
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 126 GKMDASTGKHLAVIRROPVGIYLAIPYVYNLSGSKIPALIGANVMPFPPTQSVS 185
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 116 IPSD----DRLIFTRREVPVIGATPPNPLNLSAHKILAIRATGNVYIHHPSKAPLV 171
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 186 GLVIK---APAEGLPAGVFNTITGRGSEIGDYIYEHEVNFINTGSTPVGRIKTL 241
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 172 CIELAKIENALKKKNVPLGVNLLTGAGVGEIVNEKYNMISFSGSSKVGELITRK 231
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 242 AGMRPIMLEGGKAGIYIADADIDNAKOIVAGAYVYSGQRCYAIKRVLYVEVADELA 301
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 232 AGFKKIALELGGVNNVILKQADLNKAVANALIKSFYAGOVCSIGMILVDESLDFT 291
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 302 EKISENVAKLSVGDPFNAT--VTPIVDNSADFIETSLVVDAROKAKELNFKDGRLLT 360
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 232 EMPVAKAVLANGNLDEKTDVGLISVEHAEWKEVYKAIIDEGKLLDGKRRKALFY 351
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 361 PGLFDHVTLLDKLAMEBPFPIIPIIRVKDAEEAVALANKSDPGLQSSVPTFQKAFDI 420
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 352 PTLLE--VDRDNLCTETFAPIPIRTNE--EEMIDIANSLEYGLHSAIFTNIDINKSLK 409
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 421 ANKLEVGIVHINNTKGRGPNFPFLGLKSGAGVGIYSLIEMNTNYSIVL 472
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 410 AENLEFGVINDSSLFRQDNMPFGVKKSGLGREGVYKAMEMSNIKTII 461
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 40
YELL_METUA PRELIMINARY; PRT; 470 AA.
ID YELL_METUA
AC Q97XA5;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase, NADP dependent (GapN-2) (EC
DE 1.2.1.9).
GN Name=GapN-2; OrderedlocusNames=SSO1842;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxId=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;

```

RX MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;
 RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Awaysz M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moor A., Etraugo G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL: A8006793; AAK42037.1; -.
 DR PIR: P90347; P90347.
 DR HSSP: O57693; 1KY8.
 DR GO: GO:0016491; P:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyde_dehydrog.
 DR Pfam: PF00171; Aldehyd. 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 KM Complete proteome; Oxidoreductase.
 SQ SEQUENCE 470 AA; 51603 MW; 6C22E206D5B8326C CRC64;

Query Match 31.4%; Score 752; DB 2; Length 470;
 Best Local Similarity 36.5%; Pred. No. 6.2e-37;
 Matches 172; Conservative 96; Mismatches 187; Indels 16; Gaps 7;

QY 16 SYNQIITLSPIDSSLGFPVPMTRBEVDHAKAGREALPMAALTYERAOYLHKAADIT 75
 DB 3 SLEVEIRSPNKKVIGTVKMSKDEVRGIEBEAYGFETISMPLYKTAIRKSEILR 62
 QY 76 ERDKSEIATVLAKEISKAVNASVTEVVRTADILRYAAEE-GIRLSTSDAGCKMDA---- 130
 DB 63 ERKQERLAKLAMEAKPISDSRVEVLRASRLFRQAEVGLYLE--GNTRVDAYEYP 119
 QY 131 -STGHKAVIRROPVGIYLAAPYNYPNLGSKIAIPALIGNVNMFKEPTQSGVGLV 189
 DB 120 PGNENRIVLSTREPIGVTAALPEFPNIPSPAKHVAIPALAVGSSVYVKSINPLAAYM 179
 QY 190 AARFAPAGLPAGVFNTTIGSGEIGDIYVEHEVNFINTGSTPVGORIGKLA--GMRP 246
 DB 180 KILVAGAGLPDSAVRVVGTSGSEIGDEIITHPLVGLITLTGSTGTGIALSKAISKGR- 238
 QY 247 IMLELGGKAGIYLAADLDNMAKQIYAGAYDSGCRCTAIKRVLYVEEVADELAEKISE 306
 DB 239 IIMELGSPPIILLEPANIERSIAVRAPEYAGQCNAGKRIIVROEYVDFFVAENE 298
 QY 307 NNAKLSVGDPEFNAT-VTPYIDNSADFIESTLVVDAROKAKE--LNEFRDGRLLTPGL 363
 DB 299 KAKALVGEPLDETTIVGPVINKESVENLNSVLEDAKVGGRVEILNRGPESSFFPLTM 358
 QY 364 FDIHTVLDKLAWEPPFGPIIPITRVKDAEBAVAIANKSPFGLOSVFTRDFOKAFDIANK 423
 DB 359 VIMPISDMLVLTKEVGPVIVPIVSVKSDBAIRIANSTVEGLQSAIFTDVVRALKLSRE 418
 QY 424 LEVGTTHINKTGRGPNFPFLGKSGAGVQIGRISIEAMTVKSIIVLDM 474
 DB 419 LKFGAVIINDSTRLRWDSLPEFGFKKTGIGREGVRETMLMTEINKLIAITL 469

RESULT 41
 O976X5
 ID 0976X5 PRELIMINARY; PRT; 468 AA.
 AC 0976X5;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, last annotation update)
 DE 468aa long hypothetical aldehyde dehydrogenase.
 GN OrderedocusNames-ST0064;
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=JCM 10545 / 7;
 RX MEDLINE=21456156; PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanggi M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.,
 RT "Complete genome sequence of an aerobic thermophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL: APO00981; BAB65021.1; -.
 DR HSSP: O57693; 1KY8.
 DR GO: GO:0016491; P:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyde_dehydrog.
 DR Pfam: PF00171; Aldehyd. 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 468 AA; 51257 MW; 523C477E7F0F3BE CRC64;

Query Match 31.4%; Score 751; DB 2; Length 468;
 Best Local Similarity 36.7%; Pred. No. 7.1e-37;
 Matches 171; Conservative 90; Mismatches 191; Indels 14; Gaps 6;

QY 20 ITILSPIDSSLGFPVPMTRBEVDHAKAGREALPMAALTYERAOYLHKAADITERDK 79
 DB 5 IEKPSPNKKVIGTVKMSKDEVRGIEBEAYGFETISMPLYKTAIRKSEILR 64
 QY 80 ERIATVLAKEISKAVNASVTEVVRTADILRYAAEEGIRLSTSDAGCKMDA----STGH 134
 DB 65 ERLARLAMEAKPISDKSREVEVLRASRLFRQAEBAIYLEG--KNYRVDAEYPPGENE 122
 QY 135 KLAIVIRROPVGIYLAAPYNYPNLGSKIAIPALIGNVNMFKEPTQSGVGLVIAKAPA 194
 DB 123 RIVISTREPIGVTAALPEFPNIPSPAKHVAIPALAVGSSVYVKSISTSLSIEMKLIIV 182
 QY 195 EAGLPAGVFNTTIGSGEIGDIYVEHEVNFINTGSTPVGORIGKLA--GMRPIMLEL 251
 DB 183 EAGLDSAVRIYVGTSGNEIGDEIITHPLVGLITLTGSTGTGIALSKAISKGR-1IMEL 241
 QY 252 GSKDAGIYLAADLDNMAKQIYAGAYDSGCRCTAIKRVLYVEEVADELAEKISENVATL 311
 DB 242 GSDPIIVLEADNIDRASSIAVRARYEAGQCNAGKRIIVREIYDFKVKAFKEKVKAL 301
 QY 312 SVGDPEFNAT-VTPYIDNSADFIESTLVVDAROKG--AKELNEFRDGRLLTPGLFDHYT 368
 DB 302 KYGDPULDESTDIGPINGSVEKLNKALEDAQSKGNVAVLNKGPETGFFPLSLVTNPS 361
 QY 369 LDMKLAWEPPFGPIIPITRVKDAEBAVAIANKSPFGLOSVFTRDFOKAFDIANKLEVTG 428
 DB 362 LDMVLKTEIFPIAPIVSVKSDBAIRIANSTVEGLQSAIFSDVNRALKIAKELKFGA 421
 QY 429 VHIINKTGRGPNFPFLGKSGAGVQIGRISIEAMTVKSIIVLDM 474
 DB 422 IINDSTRLRWDSLPEFGFKKTGIGREGVRETMLMTEINKLIAITL 467

RESULT 42
 O8EMK2
 ID 08EMK2 PRELIMINARY; PRT; 486 AA.
 AC 08EMK2;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, last annotation update)
 DE Aldehyde dehydrogenase (EC 1.2.-.-).
 GN OrderedocusNames=O82840;
 OS Oceanobacillus theyensis.
 OC Oceanobacillus theyensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 OX NCBI_TaxID=182710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTB31;

RP SEQUENCE FROM N.A.
 RC STRAIN-HTE831.
 RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;
 RA Takami H., Takaki Y., Uchiyama I.;
 RT "Genome sequence of *Oceanobacillus theysensis* isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments.";
 RL Nucleic Acids Res. 30:3927-3935(2002).
 DR EMBL; AP004596; BAC12817.1; -.
 DR HSSP; P05091; 1004.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyd_dehydrog.
 DR Pfam; PF00171; Aldehyd_1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 497 AA; 54140 MW; 14F219B301BE54FB CRC64;

Query Match 31.0%; Score 742; DB 2; Length 497;
 Best Local Similarity 34.8%; Pred. No. 2,6e-36;
 Matches 169; Conservative 98; Mismatches 196; Indels 22; Gaps 6;

QY 2 TBYQNYVNGEMKSSVNOQIEILSPIDSSLGFPVPMNTREVDHAKAGREALPAMALTV 61
 DB 8 TDVTKNFIDGKWDVNNNTTAVTNPANGERIYOVPLSDQSHVEAVALAKQKEMALVPA 67
 QY 62 YERAOYLAKRADIIERDKKEIATVLAKESKAVNASVTEVTRADLIRYAABEGIRLSTS 121
 DB 68 PQRAEVLAVGVGMIMCKOKERLSRLTMENKVLBARGEVQESIDMAFMAGGRLLF--- 125
 QY 122 ADEGKMDASTGKLAIVIRQYGYLAIPYNPVNLGSKAPALLIGNVNMFEPPIQ 181
 DB 126 ---GQTPPAELKDFKMSQRPVPGVGIITPMNFPIAATWKSFPALVAGNAVVMKPATE 182
 QY 182 GSVGLVLAFAEAGLPGAVNTITGRGSEIDGYVEHEVNFITGSTPGQIGTL 241
 DB 183 TPIMAYELKFIIEBAGLPGVINNVFGKSEVDSVOHDKRIVISFTSNDTGRNIASD 242
 QY 242 AG--MRPIMLELGGKAGIVLADADLDNAKQIVAGAYDSGQRCATKRVAVEEAD 299
 DB 243 CGKQKLKVLSEWGNKNAVIMDDADLDLAVEGILMAFPGSGRCACRIVVHHVKKT 302
 QY 300 LAKTISENVAKLSVGDPF--DNATVTVIDNSADFTIESLVDAKQAKELN----EFK 353
 DB 303 LBSRLAANDHLITGNGLEDGIVGPIINBAGINKIKYIEVGKQEGATLLAGTEMETN 362
 QY 354 RGRLLTPGLFDHVTILDMKLAMEEPRGPIIPITRVDAEBAVALANKSPGLQSSVFTRD 413
 DB 363 NPGYFTPTLFTDAKPMRIADSEIFGPVVSLEVSFEBALEVNNOVTFFGLSSSIFTXD 422
 QY 414 FOKAPDIANKLEVGTVHINKTGRGPDNPFLGLKSG-----AGVQIRYSIEAMTNVK 468
 DB 423 VNRVFAQGRDLDTGIVYVAAGTTGAEIHLPFGSTKGTGNGHRDSGQ----ALDVTTEWK 478
 QY 469 STYLD 473
 DB 479 AIYVD 483

RESULT 45
 027059 PRELIMINARY; PRT; 455 AA.
 ID 027059;
 AC 027059;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.
 GN OrderedLocustNames=MT978;
 OS *Methanobacterium thermoautotrophicum*.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 NCBI_TaxID=187420;

RP [1]
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Udwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of *Methanobacterium thermoautotrophicum*
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL; AE000871; AAB85474.1; -.
 DR PIR; H69230; H69230.
 DR HSSP; Q59931; 1EUH.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyd_dehydrog.
 DR Pfam; PF00171; Aldehyd_1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 455 AA; 48825 MW; A98D7BA54AA9C08B CRC64;

Query Match 31.0%; Score 741.5; DB 2; Length 455;
 Best Local Similarity 38.2%; Pred. No. 2,5e-36;
 Matches 173; Conservative 83; Mismatches 186; Indels 11; Gaps 5;

QY 9 VNGEMKSSVNOQIEILSPIDSSLGFPVPMNTREVDHAKAGREALPAMALTV 68
 DB 5 IDGEGVSGKIFTVNPFNGDEVDVPLAGRSDEVALRAHARDAMDLARKISEKL 64
 QY 69 HKAADIIEKKEIATVLAKESKAVNASVTEVTRADLIRYAABEGIRLSTSADEGCK 128
 DB 65 YVADBLKTELDPEFARLITLSEGKPIRPSRDEKVSVEFARLSAEAGL---YGESIM 121
 QY 129 DASTGKH--LAVIRQPVGIVLAIPYNPVNLGSKAPALLIGNVNMFEPPIQ 186
 DB 122 DAGIGKGLTGTVAIPLGVAIATPFNPYLMIAHKVPALAAAGNTTIKPSLEAPLSA 181
 QY 187 LVLAFAFAAGPAGVNTITGRGSEIDGYVEHEVNFINTGSTPGVQRIKLAGMRP 246
 DB 182 LKTLAMTLE--HPAAGVAVNATGRGSEVGVIIIDSPVLDKITTFGSEVGRYISARSMKK 240
 QY 247 IMLELGGKAGIVLADADLDNAKQIVAGAYDSGQRCATKRVAVEEADLAEKISE 306
 DB 241 ITLELGGNDPLVMDADIDSVAEAVRGSYISQVCAVGRMIVHEVADLAEFADKLVN 300
 QY 307 NVAKLSVGDPFDNAT--VTPVIDNSADFTIESLVDAKQAKELNEFKRDGRLTPGLFD 365
 DB 301 IINGSLRAGDPMDVTRIDVGPLINEDAIEVERVIGAAVEDGAILLCGSRGRNFVEPTVD 360
 QY 366 HTYLDMLKAMEEPRGPIIPITRVDAEBAVALANKSPGLQSSVFTRDQKAPDIANKLE 425
 DB 361 HVPQMEVIERETFGVSPPIIRFNGADEAIRANGCYALQGVTFENITRLRMARIE 420
 QY 426 VGVVHINKTGRGPDNPFLGLKSGAQVQGR 458
 DB 421 AGTIVLVNKOSTFRVDHMPF---GDGAVENWR 449

RESULT 46
 08NMBO PRELIMINARY; PRT; 484 AA.
 ID 08NMBO;
 AC 08NMBO;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE NAD-dependent aldehyde dehydrogenase (EC 1.2.1.-).
 GN OrderedLocustNames=C912668;
 OS *Corynebacterium glutamicum* (Brevibacterium flavum).
 OC Corynebacterium glutamicum (Brevibacterium flavum).

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OC NCBI_TaxId=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
 RC "Nakagawa S.;
 RT "Complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032.";
 RT Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AP005282; BAC00062.1; -.
 DR HSSP: P51977; 1BX5.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyde_dehydrog.
 DR Pfam: PF00171; Aldehyd_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KM Complete proteome; Oxidoreductase.
 SQ SEQUENCE 484 AA; 51084 MW; A2E416EE585775F CRC64;

Query Match 31.0%; Score 741.5; DB 2; Length 484;
 Best Local Similarity 36.9%; Pred. No. 2.7e-36;
 Matches 174; Conservative 93; Mismatches 184; Indels 21; Gaps 6;

QY 6 QNYVNGEW-----KSVNQIILSPIDDSLGFPVPMTRREYDHAKAGREALPAAAL 59
 DB 12 KHLIGGQWEGNSDRISTN---INPYDSVIAESKQSIADVDAYEAARKAQAEMAT 67
 QY 60 TYTTERAQLYHKAADIIRERKEIATVLAKEISKAVNASVTEVVRADLIRVAABEGIRLS 119
 DB 68 PAAESRATITVRAAELEHEHEEIVEMLRESGSTSKALETLTAGNITKESASFPGRV- 126
 QY 120 TSADGCKMDASTGKHLAVIRROPVGIYLAIPYVNVLSGSKIAPALIGGVVWFKPP 179
 DB 127 ---HGRISPSPTPKGKENVYVAVKGVGVISPMNPELNLISRSVAPALAVGNVAVIKPA 182
 QY 180 TQGSVSGLYL-AKAPAEAGLPAGVNTTGRGSEIGDYIYEHEVNFINTGSTPYGQRI 238
 DB 183 SPTPTGVGVIIPARIEEAGVPAGVISTVAGASGEIGDFVTAVPFLISFTGSTPYGRV 242
 QY 229 GKLA-----GMRPIMELGSKAGIYLAADLDLNAKQIYAGAYDVSQGCCTAIKVLVYE 294
 DB 243 GELALINGGPMKTVALELGGNAPFVVALADVDIDPAAQAAVGAFLHQGQICMSINRIVYDA 302
 QY 295 EVADLEIAEKISENVAKLSVGDPEPDNAT-VTPYIDNSADPIESLVDARQKAKELNEFK 353
 DB 303 AVHDFLEKFPYAVAKNIPTGDSASGTLVGVINQSGLKEKIELAKKEGATYQVBSR 362
 QY 354 RDGALLTPGLFPHVTLDMLAEDEBPFGPIIPRIYKDAEAAVIAANKSDFGIQQSVFTRD 413
 DB 363 IEGRLVHPHFVSDVTSDEMEIAREEIFGPIISYLKADDEHAAELANASDFGLSAAVWSKD 422
 QY 414 PQKADIANKLEVGTVHINNTKGRGPDNFPPIGLKSGGAGVGIRYSIEMT 465
 DB 423 IDRAAQFALQIDSGMVHINDLTVNDEPHVMFGSKNSGLGRNGWALIEEFT 474

RESULT 47

Q9HUR4 PRELIMINARY; PRT; 489 AA.
 AC Q9HUR4;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Probable aldehyde dehydrogenase.
 GN Ordered locus names=PA4899;
 OS *Pseudomonas aeruginosa*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxId=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardis K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004903; AAC08284.1; -.
 DR PIR: A83033; A83033.
 DR HSSP: Q63639; 1B19.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyde_dehydrog.
 DR Pfam: PF00171; Aldehyd_1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KM Complete proteome.
 SQ SEQUENCE 489 AA; 52697 MW; DECBAA8E6E73793D8 CRC64;

Query Match 31.0%; Score 741.5; DB 2; Length 489;
 Best Local Similarity 37.0%; Pred. No. 2.8e-36;
 Matches 173; Conservative 87; Mismatches 185; Indels 23; Gaps 7;

QY 11 GEMK--SSVNOIEILSPIDDSLGFPVPMTRREYDHAKAGREALPAAALTYTTERAQL 68
 DB 16 GEMRGRGRRLKYNPNPDSGLLEIBQADDIDDAIAKAAEFQPAALGPSRAAVL 75
 QY 69 HKAADIIRDYBEIATVLAKEISKAVNASVTEVVRADLIRVAABEGIRLSA---DE 124
 DB 76 YKAVEVFPDRRHEIYDWTIIRESG-----TRLKAIEIWMGAARATLTSASFPARVH 126
 QY 125 GKMDASTGKHLAVIRROPVGIYLAIPYVNVLSGSKIAPALIGGVVWFKPPQGSV 184
 DB 127 GRIVSDVPKGESRYRRAIVGVISPMNPELHLTQSIAPALAGVAVVKKPSDTPV 186
 QY 185 -SGVLAFAFAEAGLPAGVNTTGRGSEIGDYIYEHEVNFINTGSTPYGQRIKLAG 243
 DB 187 CGGLLAIRIFEEAGLPALFESVVGPSGEIDATVEHVPGLVFTSTPYGRNIGRLAS 246
 QY 244 ---NRPIMELGSKAGIYLAADLDLNAKQIYAGAYDVSQGCCTAIKVLVYEVADE 299
 DB 247 GGAHLKHALVEIGNSPVLVGLDADLEQAVNAAVFGKFLHQGQICMAINRIIVEDSLYDA 306
 QY 300 LAEKISENVAKLSVGD--PDNATYTPYIDNSADPIESLVDARQKAKELNEFKRGR 357
 DB 307 PAARVERVKGLRVGDPGRADTA-VGPIVNAKQLEGLLEKIRLAKQESAKPLYESGVGQ 365
 QY 358 LTPGLFPHVTLDMLAEDEBPFGPIIPRIYKDAEAAVIAANKSDFGIQQSVFTRDPOKA 417
 DB 366 LIAHFVFEVNTVMEIADEIFGPIVGLLRARDEHAAELANASDFGLSAAVFSRLERA 425
 QY 418 FDIANKLEVGTVHINNTKGRGPDNFPPIGLKSGGAGVGIRYSIEMT 465
 DB 426 VRFARQLRAQMTYHVDIPVNDENAAAPFGEKNSGLGRFNGWALIEEFT 473

RESULT 48

Q6M2H6 PRELIMINARY; PRT; 496 AA.
 AC Q6M2H6;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DB BENZALDEHYDE DEHYDROGENASE (EC 1.2.1.28).
 GN Name=xyIC; Ordered locus names=C92953;
 OS *Corynebacterium glutamicum* (Brevibacterium flavum).
 OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OC NCBI_TaxId=1718;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13032 / DSM 20300 / NCTB 10025;
RX PubMed=12994826; DOI=10.1016/S0168-1656(03)00154-8;
RA Kallinowski U., Bache B., Bartels D., Bischoff N., Bött M.,
RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Galglat L.,
RA Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Pühler A.,
RA Rey D.A., Rieckert C., Rupp O., Salm H., Wendisch V.F., Wiegrebe I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins";
RL J. Biotechnol. 104:5-25(2003).
DR EMBL: BX927156; CAP20692.1; -.
DR HSSP: P56533; 1A45.
DR GO: GO:0018479; F:benzaldehyde dehydrogenase (NAD+) activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002086; Aldehyde_dehydrog.
DR Pfam: PF00171; Aldehyd_1.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 496 AA; 52386 MW; C4EB32B571BE3A84 CRC64;

Query Match 31.0%; Score 741.5; DB 2; Length 496;
Best Local Similarity 36.9%; Pred. No. 2.8e-36;
Matches 174; Conservative 93; Mismatches 184; Indels 21; Gaps 6;

QY 6 QNYVNGEW-----KSSVQNIETLSPIDSSLGFPVPMTRREYVDHMKAREALPAMAA 59
DB 24 KHLIGGWEGNSDRISTN-----INPYDSVIAESQASITADVAEAAKQAAMAA 79
QY 60 TVYERAOYLKKAADIIRDKKEIATVLAKEISKAYNASVTEVVRADLRIRAAEGIRLS 119
DB 80 PAERSAIITRYAELEHEHEEIVEWLIKESGSTRKANLEITLAGNITKESASPGRV- 138
QY 120 TSADEGKMDASTGHKLAVIRROPVGIYLAIPYNYPVNLSSKIPALIGVVMFKRP 179
DB 139 ----HRISSPNTSGKENRYRYAKGVGVISPMNPLNLSIRSVAPALAVGAAVVIKPA 194
QY 180 TQGSVSGLVL-AKAFPAAGLPAGVFNITIGRSGEIGDYIVEHEVNFINTGSTPVGORI 238
DB 195 SDPRTVGVIPAIPIFEAGVPACVISTVAGAGSEIGDHPYTHAVPKLISTGSTPVGRV 254
QY 239 GKLA-----GMRPIMLEIGSGAGIVLADLDLNAKOIVAGAYVSGQRCTAIKRVLYVE 294
DB 255 GEIAINGGEMKTVALELGGNAPRVVLADADIDAQAAGAFIHOQQLCMSINRYIVDA 314
QY 295 EYVDELAEKISENVATLSVDPDPDNAT-VTPVITDINSADPIESLVVDARQKAKELNERK 353
DB 315 AVHDEFLERFVEAVVKNIPTGDPFASGTLVGVPVINDSQLKKELIAKKKEGATVQVEGP 374
QY 354 RDGRLLTTPGLFDVHTLDMKLAMEEPFGPIIPRIYVDAEAEVAIANKSDFGLOSSVFTRD 413
DB 375 IERLVHPHFVSVTYDMEIAREEITFGPLISVTKADDEAAHAAELANASDFGLSAAVWSKD 434
QY 414 FOKAFPIANKLEVGTVINNKTRGPDNPFPIGLKSGAGVQGIIRYSIEAMT 465
DB 435 IDRAAGFALQIDSGVHINDLTVNDPWHVFGSSKNSGIGRFNGDAIEEFT 486

RESULT 49
Q65NKO
ID Q65NKO PRELIMINARY; PRT; 488 AA.
AC Q65NKO;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DT YcDb (Aldehyde dehydrogenase, Aldehyde dehydrogenase).
GN Name=yCdb; ORFName=BL01646; BL100285;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxId=279010;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=15383718;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumer S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential";
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang W., de Leon A.L., Xiang H., Gustl V., Clausen I.G., Olsen P.B.,
RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
RA Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species";
RL Genome Biol. 5:R77-R77(2004).
DR EMBL: AE017333; AAU39244.1; -.
DR EMBL: CP000002; AAU21892.1; -.
SQ SEQUENCE 488 AA; 52682 MW; 14F3C8F07D2D164F CRC64;

Query Match 30.9%; Score 739.5; DB 2; Length 488;
Best Local Similarity 35.3%; Pred. No. 3.6e-36;
Matches 171; Conservative 102; Mismatches 190; Indels 21; Gaps 9;

QY 2 TREYQNYVNGEW-----KSSVQNIETLSPIDSSLGFPVPMTRREYVDHMKAREALPAMAA 58
DB 7 SKTYFNFIKRWKVAKESGMEQSLNPADTRDVLGVQKSIEDVDRAVDAKQAKAMRK 66
QY 59 LVYERAOYLKKAADIIRDKKEIATVLAKEISKAYNASVTEVVRADLRIRAAEGIRL 118
DB 67 LGAERGGFLYKADIMBQRDEIAFCATREMGKTLPEAKGEPARGIATLRVYAGGLRK 126
QY 119 STSADGKMDASTGHKLAVIRROPVGIYLAIPYNYPVNLSSKIPALIGVVMFKRP 178
DB 127 T-----GDVTPSTDSAPFYTRVPLGVGVISPMNPLNLSIRSVAPALAVGAAVVIKPA 181
QY 179 PTOGSVSGLVL-AKAFPAAGLPAGVFNITIGRSGEIGDYIVEHEVNFINTGSTPVGORI 238
DB 182 ATEIYATLCKLVISCFEEAGIPGVNNAVTPSSSGQRIAEHPDVNGITFTSSNOTGKTI 241
QY 239 GKLAGNR--PIMLEIGSGAGIVLADLDLNAKOIVAGAYVSGQRCTAIKRVLYVE 296
DB 242 GRTAFPERGAKYQLEMGKNKPVIVADADDAIVEAVISGAFSTGKCATSRVILANGV 301
QY 297 ADELAKEISENVATLSVDPDPDNAT-VTPVITDINSADPIESLVVDARQKAKELNERK 349
DB 302 YDRFKEKELQOTKEITIGSLKEDVWGMGIANKQQLDNLCTSIYAKGKGADLIFGGERL 361
QY 350 NEFK-RDGLRLTPGLFDVHTLDMKLAMEEPFGPIIPRIYVDAEAEVAIANKSDFGLOSS 408
DB 362 ADGKTENGYYIPALFDNVTSCMTIAQEEIFPVALIALADTLEELSTFANDVKGLSAS 421
QY 409 VFTRFOKAFPIANKLEVGTVINNKTRGPDNPFPIGLKSGAGVQGIIRYSIEAMT 466
DB 422 IFTQNIIRMLSTFTDEIENGLIRVNASAGVELQAFPGVQKSSSHSREG-BAKEFFTA 480

RESULT 50
Q8U2S5
ID Q8U2S5 PRELIMINARY; PRT; 506 AA.
AC Q8U2S5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase.

RC STRAIN=NCIMB 9866;
 RX MEDLINE=95014081; PubMed=7929007;
 RA Kim J., Fuller J.H., Cecchini G., McIntire W.S.;
 RT "Cloning, sequencing, and expression of the structural genes for the
 RT cytochrome and flavoprotein subunits of p-cresol methylhydroxylase
 RT from two strains of *Pseudomonas putida*.";
 RL J. Bacteriol. 176:6349-6361(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 9866;
 RX MEDLINE=20029269; PubMed=10565539;
 RA Cronin C.N., Kim J.-H., Fuller J., Zhang X.-P., McIntire W.S.;
 RT "Organization and sequences of p-hydroxybenzaldehyde dehydrogenase and
 RT other plasmid-encoded genes for early enzymes of the p-cresol
 RT degradative pathway in *Pseudomonas putida* NCIMB 9866 and 9869.";
 RL DNA Seq. 10:7-17(1999).
 DR EMBL; U96338; AAA75634.2; -
 DR FMR; T46684; T46684.
 DR HSSP; Q63639; 1B19.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyd_dehydrog.
 DR Pfam; PF00171; Aldehyd, 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW plasmid.
 SQ SEQUENCE 491 AA; 53156 MW; 70C45CE421266344 CRC64;
 Query Match 30.8%; Score 736.5; DB 2; Length 491;
 Best Local Similarity 35.8%; Pred. No. 5,5e-36;
 Matches 171; Conservative 84; Mismatches 197; Indels 25; Gaps 7;
 QY 5 YQN---YNGEWM--SSVNOIILSPIDSSLGFPVPMATREEDHAMKAGREAPMAWA 58
 DB 8 YENMSIQLAGSEWRVKGAGDLDPDPTQEKLIQIPLAREBDLDAVSARQAQVAMWA 67
 QY 59 LTYREAOYLKRAADIERDKETIATVLAKEISKANASTEVNRRADLRVAEGIRL 118
 DB 68 CGSEBAQVWLNVRIFDERRDIIMTIRESGS-----TRKAQIEWGABARAIQ 118
 QY 119 ST----SADEGKMDSTGSHKLAVIRROPVGIATAPVNYPNVLSGSKIALIGNAVY 174
 DB 119 ESMASLSRVHGRILADDPVGEKESRVYREPVGIVGITSFPMFPLHTRASLAPLALGNAC 178
 QY 175 MFKPTQGSVS-GLVLAKEPAEAGLPAVGNITTGSGEIGDYIVHEBEVNFINTGSP 233
 DB 179 VIKPABDTPTGGLLHAHFEBAGLPKGVLSVVVGSGSEIGDAFVHEVNGFISFTGSI 238
 QY 234 VGGRIKLAG---MPPIMLELGGKAGIYLADADLDNNAKQIVAGAYDYGQRCTAIR 289
 DB 239 VGNIRIRIAGGHEHLKGVALLLGNSPFPVVLADADLDQAVNMAVVGKFLQGQICMAINR 298
 QY 290 VLVEEVADELAKIESENVAKLSVGPDPMAVT-PIYDONSADFIESTLVDAKOKAXE 348
 DB 299 IYVEDSVTFEVRVRYAEVKSLSPYGDPKREYVGVYNAKKQLAGIQDKATIAKSGARV 358
 QY 349 LNEFKDGRLLTGLFDHYTLDMKLMABEPPGPIILIRKCAEBAVAIAKNSDFGLSS 408
 DB 359 MYVGEAQGVNLPHPVADVTADMEIRAEELFGPLVGIQARARBAHLLLELANSEYGLSSA 418
 QY 409 VFFRDRQKAFDIAKLEVGTVHINNTKGRQPNFPPLGLKGSAGVQGIKRIYSLEANT 465
 DB 419 VFTSLERGVKFAIRAGKTHINDIPVNDPEAPAPGEGEKSGKIGRFNGDWALIEFT 475

RESULT 53

0810X6 PRELIMINARY; PRT; 474 AA.
 AC 0810X6; Q6H242; Q6K741;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
 DE Aldehyde dehydrogenase family protein.

GN OrderedLocusNames=BA2289, BAS2135, GBAA2289;
 OS *Bacillus anthracis*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ames / isolate Porton;
 RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
 RA Read T.D., Peterson S.N., Tourasse N.J., Bailly L.W., Paulsen I.T.,
 RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
 RA Holtzapple E.K., Okestad O.A., Helgason E., Ristone J., Wu M.,
 RA Kolony J.F., Beaman M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
 RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
 RA Nelson W.C., Peterson J.D., Pop M., Kouri H.M., Radune D.,
 RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
 RA Berry K.J., Plant R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
 RA Hazen A., Cline R.T., Redmond C., Thwaitte J.E., White O.,
 RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
 RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
 RT "The genome sequence of *Bacillus anthracis* Ames and comparison to
 RT closely related bacteria."
 RL Nature 423:81-86(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ames / isolate 0581;
 RA Ravel J., Kasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
 RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
 RA Fraser C.M.;
 RT "Bacillus anthracis comparative genomics."
 RL Submitted (MAY-2004) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Stemne;
 RA Bretzin T.S., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Hitchcock P., Rubin E., Tice H.;
 RL Submitted (JAN-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AE017031; AAP26160.1; -
 DR EMBL; AE017334; AAT31413.1; -
 DR EMBL; AE017225; AAT54447.1; -
 DR HSSP; Q28399; 109J.
 DR TIGR; BA2289; -
 DR TIGR; GBAA2289; -
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyd_dehydrog.
 DR Pfam; PF00171; Aldehyd, 1.
 KW complete proteome.
 SQ SEQUENCE 474 AA; 51998 MW; 70EC5347DC9D22D CRC64;
 Query Match 30.8%; Score 736; DB 2; Length 474;
 Best Local Similarity 35.7%; Pred. No. 5,7e-36;
 Matches 168; Conservative 99; Mismatches 195; Indels 8; Gaps 5;
 QY 8 YNGEKKSVNOIILSPIDSSLGFPVPMATREEDHAMKAGREAPMAALTYVERAOY 67
 DB 6 YINGMKSVNYTKPYAPYSEBTLIAIOGTEDEVDKEAANAAMKKNNTLSAYDRAT 65
 QY 128 MDASTG--HKLAVIRROPVGIATAPVNYPNVLSGSKIALIGNVVMPKPTQGSVS 185
 DB 123 LPAAPGADGRIVATYTRKPIGVIGAITPFPPLNVHAKGVPAIAAGNTVVLAPADQTP 182
 QY 186 GLVLAKEPAEAGLPAVGNITTGSGEIGDYIVHEBEVNFINTGSPVGGRIKLAGNR 245
 DB 183 SYALIELFEBAGLPKGALNITISGPGSTVGEALVTVDVASITFTSPKGIQIKAKAGLK 242
 QY 246 PIMLEIGKDAQIVLADADL-DNAKQIVAGAYDYGQRCTAIRKLVVEEVADEIAEKI 304
 DB 243 RVTLEIGSNAAVITIDEDVELTDLIERVWKGAFVANNQVCISQVRFVHEBRDPLSL 302

```

Oy 305 SENVAKLSVGD- PNNATVTPVIDNNSAPFISLVDAOKAKEL- NEFKRDGLLTGG 362
Db 303 RKAEVTVVGDPLEETIDTSDALSISKKDVERIDINQVEAKEGANVLYGNGKRDREYFEPT 362
Oy 363 LFDHTLLMLKLAIEEPFGFILLIRIKVKADEEAVAILANKSDPGIOSVFTFRDFOKAFDIAN 422
Db 363 VLTINVPEHVSQCGEIVFGPLMTVNFPEKDEALIEQVNNRSRYGLQAGVFITNNLEFKANRAID 422
Oy 423 KLENGTVHIINNKTRGPGDNFPFLGLKGSAGVQGIIRYSIEAMTNKSVILV 472
Db 423 ELEVGVMVINIDPIPTFRVDHMPYGVKESGTGREGIKYALIEEMTEMLVCI 472

RESULT 54
Q6HJ76
Q6HJ76 PRELIMINARY; PRT; 474 AA.
AC Q6HJ76;
DT 05-JUN-2004 (TREMBLrel. 27, Created)
DT 05-JUN-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Succinic-semialdehyde dehydrogenase NMD(P)+ (EC 1.2.1.16) .
GN Name=gabD; OrderedLocustNames=Btr9727_2073;
OS Bacillus thuringiensis (subsp. konkukian) .
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=180856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27." ;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBD databases.
DR EMBL; AE017355; AAT59797.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0009013; F:succinate-semialdehyde dehydrogenase [NMD(P) . .; IEA.
DR GO; GO:0008165; P:metabolism; IEA.
DR InterPro; IPR002086; Aldelyd_dehydrog.
DR Pfam; PF00171; Aldedh; 1.
DR Complete proteome; Oxidoreductase.
SQ SEQUENCE 474 AA; 51973 MW; 9FEC4A3C76AC678F CRC64;

```

	Query Match	30.8%	Score 736;	DB 2;	Length 474;
	Best Local Similarity	35.7%	Pred. No. 5,7e-36;		
	Matches 168;	Conservative 99;	Mismatches 195;	Indels 8;	Gaps 5;
Qy	8	YVNGEWSVNOQIELSPIDSSLFQVPMRMREBYDHMKAGRELPMAWALTVERBOY	67		
Db	6	YINGDMKSVNTKPIYAPYSEETLAEIOGTEDVEVAALAAKMMKEMNTLISADRAIT	65		
Qy	68	LHKADILIERDKEELATVLAKEISKAYNASVEYVRADLIRYAABEGIRLSTSADEGK	127		
Db	66	LEKVAQKNDERREEFAELIAKEAARIPRAAREBVRTOYTFPAABEAKRT---YGETLP	122		
Qy	128	MDASTG--HKLAVIRROPVGLIALAPYNNYPVNLGSKIAPLIGNNVMEKRPPTQSVS	185		
Db	123	LDAPAGADGRIAYTIRKPIGVIIGALITPFNFPLNVAAHKGPALAGNVTVLKPADQPLS	182		
Qy	186	GLVLAKAAEAKGLPVGVEPNTTIGRGSSEIGDYIYEHBEVNFINFTGSTPYGQRIKGLAGR	245		
Db	183	SYALIELPEEEAGLPFGALNTIISGPGSTVGEALVTMDVVASITFTSPKVGIGIKAKAGLK	242		
Qy	246	PIMLELGGKDGIVLADADL--DMAKQIYAGYVDVSGQCTAIKXVLVVEEVADELAETI	304		
Db	243	RYTLELGSNAVITIDEDVELDELIERVYMGAFVANNQVCISVQRPAHEEEMDDPLSLT	302		
Qy	305	SENVAKLSVGDV--PDNATVTPVIDNSADFTISLVDAAQOKAKEL--NEFKRDGLRLTFG	362		
Db	303	RKAMEYVVGGDPLIETDVSALISKKDVERIDNWQGEANKGANLYGKNKDERVFEPT	362		
Qy	363	LFDHYTLDMKLAMEEPPGFIPLIRVKQAEVAVALANNSDFGLQSSVFTTRDFOKAFDIAN	422		

Dd		363 VLTNVEBHVSVCQGVFEGMLMTVNTFFKEDFAIEICVNNSRKGLQAGVFITNNLFKRAAID	422
Oy	423 KLEVGTVHINNKTGRGPDNPFLGLKGSGAGVGIIRYSIEAMTNVKISVL	472	
	:	: :	: : : : :
Dd	423 ELEVGVMINDIPFRVDHPMPYGVGESGTGREGIKVAIEMTEMKLVCI	472	
RESULT 55			
Q738S1			
ID	Q738S1	PRELIMINARY;	PRT; 474 AA.
AC	Q738S1;		
DT	05-JUL-2004 (TrEMBRel_27, Created)		
DT	05-JUL-2004 (TrEMBRel_27, Last sequence update)		
DT	05-JUL-2004 (TrEMBRel_27, Last annotation update)		
DE	Aldenhyd dehydrogenase family protein.		
GN	OrderedLocustNames=BCEJ2323;		
OS	Bacillus cereus (strain ATCC 10987).		
OC	Bacteria, Firmicutes, Bacillales, Bacillaceae, Bacillus.		
OX	NCB1_TaxID=222523;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	PubMed=14960714; DOI=10.1093/nar/gkh258;		
RA	Raeko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,		
RA	Shores K.A., Fouts D.B., Tourasse N.J., Angiolini S.V., Kolonay J.F.,		
RA	Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;		
RT	"The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic		
RT	adaptations and a large plasmid related to Bacillus anthracis pXO1";		
LJ	Molec Acids Res. 32:977-988(2004).		
DR	EMBL; AE017271; AAS41241.1; -.		
DR	HSSP; P20000; IAG8.		
DR	TIGR; BCEJ2323; -.		
DR	GO; GO:0016491; "Flavodoxin reductase activity; IEA.		
DR	GO; GO:0008152; "Protein biosynthesis; IEA.		
DR	InterPro; IPRO02086; Aldenhyd_dehydrog.		
DR	Fam; PF00171; Aldedn; 1.		
KW	Complete proteome.		
QC	SEQUENCE 474 AA; DCC692BB045D00AF CRC64;		

Query Match	30.7%	Score 735	DB 2	Length 474
Match Similarity	35.7%	Pred. No. 6.5e-36		
Base	Conservative 98	Misma 196	Indels 8	Gaps 5
QY	8	YVNGEWSVVAQIILSLSEIDDSLSGFPMAMREBEVDHMKAGREALPMAALITVERAOY	67	
DB	6	YINDGMSVNTKPLAYPYSEETLAEIAOQTEEPYKCAVIAKANKMEMRNLISAVDATI	65	
QY	68	LHKAAIIERKEEIALYVLAKETISKANASTEYVVRKDLIRYAAEGIRLSTADBEQK	127	
DB	66	LEKAAQKDEREEFEAEIIAKEAKPIRAAAGEVDRTQYTKFAAEAKRI--YGETLP	122	
QY	128	MDASTG--HKLAVIRROPGVIVLAAIPNYEVLNIGSKTIAPLAGANVMEKPPTOGSVS	185	
DB	123	LDAAPGADGRIRIATIRKPIGIVIGAITPNNPFLNLVAHKVGRIAAAGNVVLKPADQTPLS	182	
QY	186	GLVIAKAPAEAGLPAGVFNNTTGGSGEIGDYVEHEEVLNFTFGSTPEVGORIGKLAQMR	245	
DB	183	SYALVLEFEERAGLPRKALNITISGGSGTVGEAIVLNDVVAITTFGSSPKVGIGIKAKAGIL	242	
QY	246	PIMLELGKDGIVLADADL-DNAKQIVAAAYOYSGGRCATKRLVVEEVADELAKKI	304	
DB	243	RYTLELGSNAVITIDDELVELDELIERKMGAPFANNQVCISVQRFVHEERMBDEFLSKL	302	
QY	305	SENVAKLSVGP-FDNATVTPVLDNSADFTESLVDARQKAKEL-NEFRDGRLLTPG	362	
DB	303	QKAMETVVVGDPILLEETDVSALISKQVERIDNWQEAIVKEGASVLYGKNRDERIRFKPT	362	
QY	363	LEDVATLDMKLAMEEPGPILPIIRVYDAEEAVVLIANSDEGLQSSVTPTRFOKAFDIAN	422	
DB	363	VLTNVPEEYVSVQCEVGRPLMTVTTFKEFDALIEQVNSHRIQLOAGVTTNMLFKMRAID	422	
QY	423	KLVEGVTHINNKTRGPDNPFPLQKSGSAGVQGIKRSIEMATNVKSLIVL	472	
DB	423	ELFEGVAVINDIPTEFVDMHPYGVGKESGTREGIKYALIEMTENKVLVCI	472	

RESULT 56
 09KC36 PRELIMINARY; PRT; 482 AA.
 ID 09KC36
 AC 09KC36;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-MAR-2004 (TREMBLrel. 26, last annotation update)
 DE NADP-dependent glyceralddehyde-3-phosphate dehydrogenase.
 GN OrderedLocustNames=BH1738;
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125;
 EX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 halodurans and genomic sequence comparison with Bacillus subtilis.",
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001513; BAB05457.1; .
 DR PIR; B83867; B83867.
 DR HSSP; Q59931; 1EUN.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyd_dehydrog.
 DR Pfam; PF00171; Aldehyd; 1.
 KW Complete proteome.
 SQ SEQUENCE 482 AA; 52424 MW; 29DFBCABEB1F8D4D CRC64;
 Query Match 30.7%; Score 735; DB 2; Length 482;
 Best Local Similarity 35.0%; Pred. No. 6.7e-36;
 Matches 164; Conservative 90; Mismatches 209; Indels 6; Gaps 3
 QY 9 VNGEWKSSVNOQIEILSPIDSSLSGLGVPAMTREEDVHAKAGREALPAAALTVYERAOYL 68
 DB 12 ISDQOHQOQSIEVIDPCTNHVIAIVPATEADAKAIEAAKGAEIRAROMPHERRAIL 71
 QY 69 HKAADIIRDKEEITVLAKESIKAYNASVTEVVRATDLIRAAEEGRILSTSDGGM 128
 DB 72 KRTAEIVTRDELFPAQTLAKESGSKTIRERSEVRSCIEFTLTSGEARLHGETIPFSOM 131
 QY 129 DASTGH--KLAVIRROPVGIYLAIPVYVPAVLSSGSKTAPALIGNVVMFKPPTGOSVSG 186
 DB 132 ---PCHERRVGVGFYFPFGIIVAIIPFPDPLNLVHKIGPALAAGNSLIRKSSPTPLSA 188
 QY 187 LVYAKAFPAEAGIPAGVFNFTTGRGSEIGDYVEHEEVNFINTGSTPVGQRIKGLAKMRP 246
 DB 189 LRUVVATLEAGIPKKTIVGVITGHSVIGPTTLKHDVRLISPTGGEYEGEKIARSGAYKK 248
 QY 247 IMLEIGKADAGIVLADADLDLNAKQIVGAYDVSQRCATIRVIVVEBVDLELAETISE 306
 DB 249 LAMEIGNSPTIVLDDAEIMEAVASCVSGAFACAGQNCIGVQRITVYGOSVFNSFIQSYVA 308
 QY 307 NVAKLSVGDPEPDNAI-VTPVIDNSADEIESIIVADAROKAGAKELNEFKRDGLLTPGLFD 365
 DB 309 QTKQLRLDKQSEEDFDIGPMISEKAKRIERWVEAKBEGARVLITGSRITGAYFEPVILT 368
 QY 366 HTYLDMLKAMEEPPFPILPIIRYKDAEAVAIANKSDGLQSGVSVTRPQKAFDLANTLE 425
 DB 369 NVSPESRLAKEEAFPIYIIGSVHSLTEIYAAADVDFGLQGLGELTNLNLTAFAISAIETLE 428
 QY 426 VGTVAHNKTKRGPDNPFLGLKSGAGVQGIYRSIEAMTVKYSIVLDM 474
 DB 429 VGGIWNDSDVIRIDAMPFGGIIKSGLGREGVRVAIEEMTEQKVAAFHL 477

ID	Q8FWC7	PRELIMINARY;	PRT;	471 AA.
AC	Q8FWC7			
DT	01-JUN-2002 (TrEMBLrel. 21, Created)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	NAD-dependent aldehyde dehydrogenase.			
GN	OrderedLocustNames=MK1108;			
OS	Methanopyrus kandleri.			
OC	Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;			
OC	Methanopyrus.			
OK	NCBI_TaxID=2320;			
OK	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=AV19 / DSM 6324 / JCM 9639;			
RX	MEDLINE=21927647; PubMed=11930014; DOI=10.1073/pnas.032671499;			
RA	Slesarev A.I., Mezhvaya K.V., Makarova K.S., Polushin N.N.,			
RA	Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,			
RA	Natale D.A., Rogozin I.B., Tatunov R.L., Wolf Y.I., Stetter K.O.,			
RA	Mal'kh A.G., Koonin E.V., Kozlovskiy S.A.,			
RT	"The complete genome of hyperthermophilic Methanopyrus kandleri AV19			
RT	and monophyly of archaeal methanogens."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).			
DR	EMBL; AB010400; AAC02321.1; .			
DR	HSSP; P51977; IBSX.			
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	Pfam; PF00171; Aldedh; 1.			
DR	PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKOWN_1.			
KW	Complete proteome.			
SQ	SEQUENCE 471 AA; 51147 MW; E3695386CECD7B3D CRC64;			
Query Match	30.7%; Score 734; DB 2; Length 471;			
Best local Similarity	36.7%; Pred. No. 7,4e-36;			
Matches 173; Conservative	90; Mismatches 201; Indels 8; Gaps 5;			
QY	4 EYQNVNNGEKKSSNQNEILSPIDDSLGVPVMTREBYDHAKKAGELAPAAALTYE 63			
DB	2 EYGLIGGEWLEGDREIVENPYDGSVGVLA-PEYDVALLKDAREGRRRREPTYE 60			
QY	64 RAQYIHLRAADIIERDEKEIATVLAKEISKAYNASVTEVNTADILIRAAEEGRLTSAD 123			
DB	61 VREHLAPAAHLILKHEBELAELIALBEGKRIIDARAYVYTRREYLRISAABERLYGETL 120			
QY	124 EGGKMDASTGH--KLAVIRPQVGIYLAIPYVNVLSGSKIALPALIGNVVMFKPPTQ 181			
DB	121 PG--DAQRGRTEALLTVREPVGVLSIPPNPFLILPHKLGLPALAACSVVHKPATV 177			
QY	182 GSVSGVLTAAPFAEAGIPAGVNTITRGSEIGDYIYHEEUNFINFTGSTPVGQRGK 241			
DB	178 TPISLSILAEIILDAGVEPLALOVVPGALAEBELA-RADFPALSTGSRSGVEHIREI 236			
QY	242 AGMPPIMLELGGKAGVIVLADLDNNAKQIVGAYDYSGQRCTAIKRVLVVEEVADELA 301			
DB	237 SPIRITILEGGANDPVIVDETADVEAAEAALVAGACYHAGQVCIAVERAIVVEDYEEFL 296			
QY	302 EKISENVAKLSVGPDPDNAT-VTPVIDDNSADPIESLVVDARQKGAKEINEFKRDGRLT 360			
DB	297 EAANEVAEASLKVGDPDLEDOTDVGRLDDGAVEKVRHRYVEDAVRGAEILILGGEPEGRFLP 356			
QY	361 PGLFDHYTLDMKLAMEEPGPILPITIRVKDAEEAVALANKSDPGLQSSVFTTRDPQKAFDI 420			
DB	357 PTVIADVPEDALVARRETFGPVIVIRAKDFEBAIRIANSTDGLNAHVAFTERLIDRAVKA 416			
QY	421 ANKLEVEVTHIINKTGSGRGNPFLILGKSGGAAVVGIRVSEIEMTAVKSVVL 472			
DB	417 TRKLEAGVIVNBSSTIYRADYMFEGGVKASGVREGVQAVREFTEKTIYVI 468			
RESULT 58				
063B07				
ID	063B07	PRELIMINARY;	PRT;	474 AA.
AC	063B07			
DT	25-OCT-2004 (TrEMBLrel. 28, Created)			

DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, last annotation update)
 DE Succinic-semialdehyde dehydrogenase NAD(P) + (EC 1.2.1.16).
 GN Name=gabd; ORNames=BTZK2069;
 OS Bacillus cereus ZK.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=288681;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ZK;
 RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RA "Complete genome sequence of *Bacillus cereus* ZK."
 RT Submitted (Jul-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL: CP000001; AU18186.1; -.
 KM Oxioreductase;
 SQ SEQUENCE 474 AA; 51985 MW; 24E6863CBA616C48 CRC64;

Query Match 30.6%; Score 732; DB 2; Length 474;
 Best Local Similarity 35.7%; Pred. No. 9, 9e-36;
 Matches 168; Conservative 98; Mismatches 196; Indels 8; Gaps 5;

QY 8 YNNGEWSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAMALTYERACQ 67
 DB 6 YINGWKSQVNTYKPLTAPYSEETLAEIAOQTEEDVKEAVALAKMKEMNTLSAYDRATI 65
 QY 68 LHKADIIERDKKEIATVAKESISKAYNVSVTEVNTALINVAEGIRLSTSDSEGK 127
 DB 66 LKKVAKMDERREBEFEITAKAKAKIRARGEVDVTOTYFPALEAKRI--YGETLP 122
 QY 128 MDASTG--HKLAVIRQPVGIYLAIPYNYPVNLSSGKIAPALLIGNVVMFKPPTQGSVS 185
 DB 123 LDAAGADGRIVYTRKPIGVIGAITPMPPLNLVHKGPALAGNTVVLPAQOTPLS 182
 QY 186 GLVLAKAFKAEAGLPAGVNTTTRGSEICDYIYEHEVNFINTGSTVGORIGKLAKR 245
 DB 183 SYALLLELFEAEGLPKGALNIIISGPGSTGEALVTNDVASITFTGSPKIGIKAKAGK 242
 QY 246 PIMLELGGADAGIVLADNL-DNAKQIYAGAVDSGQCTAKIRLVYVEVADLAELKI 304
 DB 243 RYTLBGSNAAYIIDEVDELTELERVKGAFVNNQVCISQVRVHEERNEFLSTL 302
 QY 305 SENVAKLISYGD-PDNATVTPYIDNSADPISLVVDARQKAKEL-NFKRDGRLLPFG 362
 DB 303 RRAMETVVVGDPLEIETDVSALISKDVRIDNVOGEALVEGANLVYGGKRDVRFET 362
 QY 363 LFDHVTLDKLAWEPPGPIPIIRYKAEAEVAVALANKSPGLQSSVFTRDFOKAFDIAN 422
 DB 363 VLTNPBEHVSVOQEVFGPLMTVNTFKEFNEALIEQVNNSRYGLOAGVFTNNLFFKMRALD 422
 QY 423 KLENGTVHNNKTRGCPDNFPPLGLKGSAGVQGIKYSIEANTNYSITL 472
 DB 423 ELEVGQVMINDIPTRFVDHMPYGVKESGTGRBGKRYALEEMTEKLVC 472

RESULT 59
 GABD_ECOLI
 ID GABD_ECOLI STANDARD; PRT; 482 AA.
 AC P25526; P78207; P78208; P78209;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, last sequence update)
 DT 25-JAN-2005 (Rel. 46, last annotation update)
 DE Succinate-semialdehyde dehydrogenase [NADP+] (EC 1.2.1.16) (SSDH).
 GN Name=gabd; OrderedLocustName=b2661;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KL2 / JMI03;
 RX MEDLINE=94127927; PubMed=8297211;

RA Niegemann E., Schulz A., Bartsch K.;
 RT "Molecular organization of the *Escherichia coli* gabd cluster:
 RT succinylate sequence of the structural genes gabd and gabb and
 RT expression of the GABA permease gene.";
 RL Arch. Microbiol. 160:454-460 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KL2 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KL2;
 RX MEDLINE=97349980; PubMed=9205837;
 RA Yamamoto Y., Alba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
 RA Mizubuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
 RA Oshima T., Oyama S., Saito N., Sato Y., Satoh Y., Sivasubraman S.,
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yamagata S., Horiuchi T.;
 RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*-
 RT K12 genome corresponding to 50.0-68.8 min on the linkage map and
 RT analysis of its sequence features.";
 RL DNA Res. 4:91-113 (1997).
 CC -1- CATALYTIC ACTIVITY: Succinate semialdehyde + NAD(P) (+) + H(2)O =
 CC succinate + NAD(P)H.
 CC -1- PATHWAY: 4-aminobutyrate (GABA) degradation.
 CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 CC -1- CAUTION: Ref. 3 sequence differs from that shown due to frameshifts
 CC in positions 51 and 190.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC CC
 DR EMBL: M88334; AAC36831.1; -.
 DR EMBL: U00096; AAC75708.1; -.
 DR EMBL: D90890; BAA16522.1; ALT_FRAME.
 DR EMBL: D90890; BAA16523.1; ALT_FRAME.
 DR EMBL: D90890; BAA16524.1; ALT_FRAME.
 DR PIR: F65045; F65045.
 DR HSSP: Q28399; 109J.
 DR ECHOBASE; EB1305; -.
 DR Ecogene; EG11329; gabd.
 DR InterPro; IPR02086; Aldehyde_dehydr.
 DR Pfam; PF00171; Aldehyd_1.
 DR TIGRFAMs; TIGR01804; BADH_1.
 DR TIGRFAMs; TIGR01722; MMSDH_1.
 DR TIGRFAMs; TIGR01780; SSADH_1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS_1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU_1.
 DR Complete proteome; NADP; Oxioreductase.
 FT NP BIND 233 238 NAD(P) (ADP part) (By similarity).
 FT ACT_SITE 255 255 By similarity.
 FT ACT_SITE 289 289 By similarity.
 SQ SEQUENCE 482 AA; 51720 MW; 091538F8741DB0CF CRC64;

Query Match 30.3%; Score 726; DB 1; Length 482;
 Best Local Similarity 34.3%; Pred. NO. 2, 3e-35;
 Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;

QY 6 QNYVNGEWSVNOIEILSPIDSSLGFPVPMTRBEVDHANKAGREALPAMALTYE 63
 DB 12 QALINGEWLDANNGRAIDVTNPANGDKLGSVPKMGADETRAALDANRALPAMRALTAKE 71

```
QY 64 BAQYLHAKADIIERDKERIATVLAKEISKAYNAVTEVTRADLIRYAAREGIRLSTSD 123
DB 72 RATTILNENLWMEHODDLARLMTLLEOGKPLAEKAEISYAASFIEWAEBEGRIYDPT 131
QY 124 ECGKMDASTGTHKLAIVRRQPGVIVLAIPYNYPNVLSGSKIAPALIGNVVMPFPPTQGS 183
DB 132 PGHQAD-----KRLIYIKOPISGTAALTTPMNFPAAMITRKAGPALAACCTMVLKPSQTP 186
QY 184 VSGDLVLAKEAFAEAGLPAVFNITIGRSGEIGDYIVHEEVENFINFTGSPVQORIGKLAG 243
DB 187 FSAALAEIALIRAGVAGVFNVTGAGAVGNELTGNPLVYRKLSFGTSGTEIGRLMEQCA 246
QY 244 --MRPIMLELGGKAGIYVLADADLDMAAQIVAGAYDSGQRCRTAIRVLVVEEVDLA 301
DB 247 KDIKRYKSLLEGAPPIVFDDADLDKAVGALASKFRMAGQTCVCANRLTYVDGVDYDRA 306
QY 302 EKISEVNAKLSVGPDPDN--ATVTPVIDNSADPIESLVVDARQKAKEL---NEFRDGR 357
DB 307 EKLQCAVSKLHIDGIDNGVTTIGPLIDEKAVAKVEHIALDEKARVCGGKAHERGGA 366
QY 358 LLTPGLFDHYTLDMKLAWEPEPGPILPIIRVKDAEBAVAIANKSDFGLOSSVFTDPOKA 417
DB 367 FPGPITLVDPNAKYSKEETFGPLAPLFRPKDEADVIAQANDTERGLAAVFAARDLSRV 426
QY 418 FDIANKLEEVGTVINNKTKG--RGPDPNPFGLKSGAGVQIGIRYSIEAMTNVKSIVL 474
DB 427 FRVGALEXYGIVGIN--TGIIISNEVAPFGGIRKASGIGREGSKYGIEDYLEIKYMCIGL 482
```

RESULT 60

```
Q72100 PRELIMINARY; PRT; 488 AA.
ID 072100
AC 072100;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Succinate-semialdehyde dehydrogenase (EC 1.2.1.16).
GN Name:sgad; OrderedlocusNames=LMOI2365_0935;
OS Listeria monocytogenes (serotype 4b / strain F2365).
OC Bacteria; Firmicutes; Bacilliales; Listeriaceae; Listeria.
OX NCBI_Taxid=265669;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15115801; DOI=10.1093/nar/ghs562;
RA Nelson K.E., Foute D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA Kolonay J.F., Rasko D.A., Argoudin S.V., Gill S.R., Paulsen I.T.,
RA Peterson J.D., White O., Nelson W.C., Nieman W.C., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Forberger H.A., Tran B., Kacharion S., Wondolting L.D., Unlrich G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
RT borne pathogen Listeria monocytogenes reveal new insights into the
RT core genome components of this species.";
RL Nucleic Acids Res. 32:2386-2395(2004).
DR EMBL; AE017325; AAT03714.1; -.
DR HSSP; P20000; 1A68.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0009013; F:succinate-semialdehyde dehydrogenase [NAD(P) .]; IEA.
DR GO; GO:0008152; F:metabolism; IEA.
DR InterPro; IPR002086; ALDEHYD_DEHYDR.
DR Pfam; PF00171; Aldehdh; 1.
KW Complete proteome.
SQ SEQUENCE 488 AA; 53194 MW; ACE7841259C0D152 CRC64;
```

```
Query Match 30.3%; Score 726; DB 2; Length 488;
Best Local Similarity 34.2%; Pred. No. 2.3e-35;
Matches 163; Conservative 95; Mismatches 200; Indels 18; Gaps 6;
QY 8 YVNGEKSSVNO--EILSPIDSSLGFPAMTRBEVDHAMKAGREALPAMALTYVERA 65
DB 16 FINGKKTDDNKETKDIVAPNAGEVYAKIAQAGPNETTKAIRAKAEPDMAKMELADRV 75
```

```
QY 66 QYLIHRAADIIERDKERIATVLAKEISKAYNAVTEVTRADLIRYAAREGIRLSTSDAG 125
DB 76 KLIHAKIADMEKBAQTLAKIMTLLEOGKPLKESKEGYLTGAENFRFAAEARRL-----YG 130
QY 126 GMDASTGTHKLAIVRRQPGVIVLAIPYNYPNVLSGSKIAPALIGNVVMPFPPTQGS 185
DB 131 ETIPAPNNHAF--IVKKQPIGVAAITPMNFGGAMTRKLAIPALATGNTIVLPGSGDTPLS 189
QY 186 GLVLAKEAFAEAGLPAVFNITIGRSGEIGDYIVHEEVENFINFTGSPVQORIGKLAG-- 243
DB 190 ALAIEFIEFEALPGVANNITWVGSSKEIGETLITSDQVAKLFTGSKYQGLTFQSAFT 249
QY 244 MRPIMLELGGKAGIYVLADADLDMAAQIVAGAYDSGQRCRTAIRVLVVEEVDLA 303
DB 250 LKKSILELGGHAPPIVFDDANLDAVNDLVAAKFRNNGCVCSPNKIPAKKIKKFTYA 309
QY 304 ISEVNAKLSVGPDPNATVTPVIDNSADPIESLVVDARQKAKELNEFK-----DG 356
DB 310 LVAKBQLKVGGLDVNVGPIREDADLDKIDKLNATDKAKAVLTGGGRLTGSDDYDK 369
QY 357 RLITPGLFDHYTLDMKLAWEPEPGPILPIIRVKDAEBAVAIANKSDFGLOSSVFTDPOKA 416
DB 370 NRYKPTLVDPNATVTPVIDNSADPIESLVVDARQKAKELNEFK-----DG 429
QY 417 AFDIANKLEEVGTVINNKTKG--RGPDPNPFGLKSGAGVQIGIRYSIEAMTNVKSIVL 472
DB 430 VEKVGALEXYGIVGIN--TGIIISNEVAPFGGIRKASGIGREGSKYGIEDYLEIKYMCIGL 484
```

RESULT 61

```
Q8FUE1 PRELIMINARY; PRT; 499 AA.
ID 08FUE1
AC 08FUE1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Putative aldehyde dehydrogenase.
GN OrderedlocusNames=CE0079;
OS Corynebacterium efficiens.
OC Corynebacterium; Actinobacteria; Actinobacteriales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_Taxid=152794;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens ";
RL Genome Res. 13:1572-1579(2003).
DR EMBL; AP005214; BAC16889.1; -.
DR HSSP; P56533; 1A4S.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; F:metabolism; IEA.
DR InterPro; IPR02086; ALDEHYD_DEHYDR.
DR Pfam; PF00171; Aldehdh; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 499 AA; 53061 MW; F642B86E33C1BA8F CRC64;
```

```
Query Match 30.3%; Score 725; DB 2; Length 499;
Best Local Similarity 35.4%; Pred. No. 2.8e-35;
Matches 166; Conservative 102; Mismatches 185; Indels 16; Gaps 8;
QY 6 QNYVNGEM--KSSVQIIEILSPIDSSLGFPAMTRBEVDHAMKAGREALPAMALTYVE 63
DB 22 QIFING--VWDGISESTTIANPFNGEAEVATIRMSQEDVDYAKATATAAEVQEMAAVPEVK 80
```

```
QY 64 RAQYHAKADITIERDKEBIATVLAKETI-SKAYNAVTEVTRADILRYAAEEGIRLSTSA 122
Db 81 KARIINKAAQVIEDRDALVELIIRRESGSTALKANIEGLAMASL-REAAETPARIT--- 136
QY 123 DEGGHMDASTGHKLAVIRROPVGIYALAPVNYPVNLSSGSKTAPALIGNVVMFKPPFQ 182
Db 137 -GQLPSTPEKSNVYFREPGLVGVLSFPMNPPALSKRSVAAPALACNGVLPASDT 194
QY 183 S-VSGIYLAKAFAEAGLPAGVFNTITGRSEIGDYIYEHEEVNFINTGSTPVGQIRGL 241
Db 195 PIVGSTILLAHIFEAAGLPFGGLFNVVAGSEIGDYFVEHRIIRLLISFTGSTVGGQVAT 254
QY 242 AG---MRPIMELGSKDAQIVLADLDLNNAAKOIVAGAYDVSQRCIAIKVLYVEEVA 297
Db 255 ATGCKYIKKVALLELGSNAPMVVLSIDADIKKVAALASMSGFLHQGICMAVNRITVEAPY 314
QY 298 DELAKISIRNVAKLSVGDPPDNAT-VTPYIDNSADFIESLVVDAROKAKELNEFKDG 356
Db 315 DQFLGFAFARVSKVAYGQDLDATWGPVINDSQVESVSGKIDLABEGARAVSGPIEG 374
QY 357 RLITFGLFDHVTLDKMLAEEBFGPILPIIRVKDAEVAVALANKSDPGLQSSVFTRDFOK 416
Db 375 RYIAHVFADVTPEMELFREBELFGVAVGVKADNEHLELANDPREGLSAVFTEDLOR 434
QY 417 AFDIANKLEVGTVHINNKTGRGPDNFPFLGLKSGAGVQIRYSIEAMT 465
Db 435 GINFARVBEAGMTHTINDITVNDENAHVMFGEGKNSGIGRENGEMALDEFT 483
```

RESULT 62

```
Q73EK2 PRELIMINARY; PRT; 483 AA.
AC 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Succinate-semialdehyde dehydrogenase (NADP+) (EC 1.2.1.16).
GN Name=gabd; OrderedlocusNames=BCE0356;
OS Bacillus cereus (strain ATCC 10987);
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Raeko D.A., Ravel J., Oakstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angluoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RT adaptations and a large plasmid related to Bacillus anthracis pXOI";
RL Nucleic Acids Res. 32:977-988(2004).
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL; AB01245; AAS39292.1; -.
DR HSSP; P20000; IAG8.
DR TIGR; BCE0356; -.
DR GO; GO:0004028; F:aldehyde dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0009013; F:succinate-semialdehyde dehydrogenase [NAD(P)-]; IEA.
DR GO; GO:0008153; P:metabolism; IEA.
DR InterPro; IPR02086; Aldehyd_dehydrtog.
DR InterPro; IPR010102; SSADH.
DR Pfam; PF00171; Aldehd; 1.
DR TIGRFAMs; TIGR01780; SSADH; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR Complete proteome; Oxidoreductase_
SQ SEQUENCE 483 AA; 52276 MW; AIDBFA7E035D42E6 CR664;
```

```
Query Match 30.2%; Score 722.5; DB 2; Length 483;
Best Local Similarity 35.5%; Pred. No. 3.7e-35;
Matches 170; Conservative 84; Mismatches 202; Indels 23; Gaps 8;
```

```
QY 8 YNNGWKSSVNOIETILSPDSSLGFPVPMTRBVDHANKAGREALPMAALTITVRAQY 67
Db 16 YINGEWITLQEQIEVNNPATKEIFATVPKGVTEAKQAVDAHAHEAFKWSKLTADRAVK 75
```

```
QY 68 LHKADITIERDKEBIATVLAKETISKAYNAVTEVTRADILRYAAEEGIRLSTSADEGK 127
Db 76 LKNMFTLIDENKEEIAIMTREOGKFPABALGEVNVNANFPEWVAEBGKV-----YGM 130
QY 128 MDASTGHKLAVIRROPVGIYALAPVNYPVNLSSGSKTAPALIGNVVMFKPPFGSUGL 187
Db 131 IPASHPNRRIIVMKOPGVMAAITPWNFPAMITRKVAIPALACGTAVVKPASQTPFAL 190
QY 188 VLAKAFAEAGLPAGVFNTITGRSEIGDYIYEHEEVNFINTGSTPVGQIRGLAG--MR 245
Db 191 KLAELAEADLPKGINVITGSAPALADTWEDGRVKVSTGSTGELGKELMASHAQTMK 250
QY 246 PIMELGSKDAQIVLADLDLNNAAKOIVAGAYDVSQRCIAIKVLYVEEVADEIAEKIS 305
Db 251 KVSLELGSHAPFIWNDDLDKXVBAVIGSKFRNAGQCICITNNYFVQEEYEVFAEFQ 310
QY 306 ENVAKLSVGDPP-DNATTPYIDNSADFIESLVVDAROKAKELNEFKR---DGRLLT 360
Db 311 KSVGQLKVGDFGQGTGVPILIDENAVSKVQEHADAIQKGTLYYGQKYALDGHFMQ 370
QY 361 P---GLPHTVTLDMKLAEEBFGPILPIIRVKDAEVAVALANKSDPGLQSSVFTRDFOK 417
Db 371 PTVGLADNTMLCMN---BETFGVAPFAKTYEBEYERANMTPYGLAAVITFKDISQA 427
QY 418 FDIANKLEVGTVHINNKTGRGPD--NPPFLGLKSGAGVQIRYSIEAMTVKSIYLDM 474
Db 428 FOISEALEYGIIGLNDGL---PSVAQAPFGGFKSGIGREGHGHIIEYLEIKTISGL 483
```

RESULT 63

```
Q81ZE2 PRELIMINARY; PRT; 483 AA.
AC 081ZE2; 061488; 06KXZ4;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Succinate-semialdehyde dehydrogenase (NADP+).
GN Name=gabd; OrderedlocusNames=BA0327;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porten;
RX MEDLIN=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Oakstad O.A., Helgason E., Ristone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nietman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related Bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Frazer C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (May-2004) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sterner;
RA Bretin T.S., Bruce D., Chailcombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longhure J., Lucas S., Okinaka R.,
```

RA Richardson P., Rubin E., Tice H.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 DR EMBL: AE017025; AF24360.1; -
 DR EMBL: AE017334; AF29418.2; -
 DR EMBL: AE017225; AF252643.1; -
 DR HSSP: P51977; 1BX5.
 DR TIGR: BA0327; 1BX5.
 DR TIGR: GBA0327; -
 DR GO: GO:0004028; F:aldehyde dehydrogenase activity; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehydrog.
 DR InterPro: IPR010102; SSADH.
 DR Pfam: PF00171; Aldehyd; 1.
 DR TIGRfam: TIGR01780; SSADH; 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 DR Complete proteome; Oxidoreductase.
 SQ SEQUENCE 483 AA; 52258 MW; 04C6B3D3F648F277 CRC64;

Query Match 30.2%; Score 721.5; DB 2; Length 483;
 Best Local Similarity 35.7%; Pred. No. 4.3e-35;
 Matches 171; Conservative 82; Mismatches 203; Indels 23; Gaps 8;

QY 8 YVNGEKSSVNOIEILSPIDSSIGFVPAMTREEDVHAKAGREALPANAALTVERAQQ 67
 16 YINGEWITTOEOEVNPNPATKEIFATVPKGGVTEAKQVDAHAHEAFKSKLTAADRAK 75
 68 LHKADIIERDKEEIAITVLAKETSKAYNASVTEVTRADLIRYAAEEGIRLSTSADEGK 127
 76 LKKWFTLIDENKEEIAIMTKEGKPPAALGEVYNANSFVEVYAEGRV-----YGEH 130
 DB 128 MDASTGHKLAVIRQVGVILAIAPYNYPNVLSGSKIAPALIGNVMMFKRPTQGSVGL 187
 131 IPASHPNKRILVWKQVGVAAITPWNFPAMITRRKAPALAGCTAVVAPASQTLPLAL 190
 QY 128 MDASTGHKLAVIRQVGVILAIAPYNYPNVLSGSKIAPALIGNVMMFKRPTQGSVGL 187
 131 IPASHPNKRILVWKQVGVAAITPWNFPAMITRRKAPALAGCTAVVAPASQTLPLAL 190
 DB 128 MDASTGHKLAVIRQVGVILAIAPYNYPNVLSGSKIAPALIGNVMMFKRPTQGSVGL 187
 131 IPASHPNKRILVWKQVGVAAITPWNFPAMITRRKAPALAGCTAVVAPASQTLPLAL 190
 QY 188 VLAKAPAEAGLPAGVFNITGRSEIGDYIVHEEVNFINTGSTPVGQRIGLAG--NR 245
 191 KLAEHLAHEADIPRGVINITYGSAKALADYMEDEGRKXVSFTSTIGKELMASAQTMK 250
 DB 246 PIMLEIGKDGAGIVLADADLDNNAKOIVAGAYDSGORCTAIRVLVEEVADELAEKIS 305
 251 KVSLEIGGHAFFIVMNDADLDKAVEAVIGSKPRNAQOTCTCTRVPVQEEVYAPVEKFE 310
 QY 306 ENYAKLSVGDPF--DNATVTPVIDNSADFIESLVDAARQKAKELNEFRK----DGRLLT 360
 311 KAVGQKLKVDGFGDGTTPVPLIDENAVSKVQEHIBDAIQGQVLYGQGVAAELDGHFMQ 370
 DB 361 P---GLFDHTVTLDMKLAMEEPFGPIPIIRVDAEEVAVALANKSDPGLQSSVFTTRDFOKA 417
 371 PVIYGLANDTMLCMN---EETFGPVAPVAKFKTVEEIVIRAHNTPPGLAAVIFTKDISQA 427
 QY 418 FDIANKLEVGTVHINKTKRGPD--NPPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDM 474
 428 FOISEALEYGIIGLNDGL---PSVAQAPFGGFGESGIGREGGHGIEEYLEIKYISLGL 483

RESULT 64
 Q6HP75
 ID O6HP75 PRELIMINARY; PRT; 483 AA.
 AC O6HP75;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 GN Name=aldehyde dehydrogenase (NAD(P)+) (EC 1.2.1.16).
 DN Name=aldehyde dehydrogenase (NAD(P)+) (EC 1.2.1.16).
 OS Bacillus thuringiensis (subsp. konkukian).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=180856;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=97-27;
 RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hall K.,

RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.;
 RL "Complete genome sequence of *Bacillus thuringiensis* 97-27.";
 CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 DR EMBL: AE017355; AF261348.1; -
 DR GO: GO:0004028; F:aldehyde dehydrogenase activity; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0009013; F:succinate-semialdehyde dehydrogenase [NAD(P) . .]; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002086; Aldehyd_dehydrog.
 DR InterPro: IPR010102; SSADH.
 DR Pfam: PF00171; Aldehyd; 1.
 DR TIGRfam: TIGR01780; SSADH; 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 DR Complete proteome; Oxidoreductase.
 SQ SEQUENCE 483 AA; 52258 MW; 04C6B3D3F648F277 CRC64;

Query Match 30.2%; Score 721.5; DB 2; Length 483;
 Best Local Similarity 35.7%; Pred. No. 4.3e-35;
 Matches 171; Conservative 82; Mismatches 203; Indels 23; Gaps 8;

QY 8 YVNGEKSSVNOIEILSPIDSSIGFVPAMTREEDVHAKAGREALPANAALTVERAQQ 67
 16 YINGEWITTOEOEVNPNPATKEIFATVPKGGVTEAKQVDAHAHEAFKSKLTAADRAK 75
 68 LHKADIIERDKEEIAITVLAKETSKAYNASVTEVTRADLIRYAAEEGIRLSTSADEGK 127
 76 LKKWFTLIDENKEEIAIMTKEGKPPAALGEVYNANSFVEVYAEGRV-----YGEH 130
 DB 128 MDASTGHKLAVIRQVGVILAIAPYNYPNVLSGSKIAPALIGNVMMFKRPTQGSVGL 187
 131 IPASHPNKRILVWKQVGVAAITPWNFPAMITRRKAPALAGCTAVVAPASQTLPLAL 190
 QY 128 MDASTGHKLAVIRQVGVILAIAPYNYPNVLSGSKIAPALIGNVMMFKRPTQGSVGL 187
 131 IPASHPNKRILVWKQVGVAAITPWNFPAMITRRKAPALAGCTAVVAPASQTLPLAL 190
 DB 128 MDASTGHKLAVIRQVGVILAIAPYNYPNVLSGSKIAPALIGNVMMFKRPTQGSVGL 187
 131 IPASHPNKRILVWKQVGVAAITPWNFPAMITRRKAPALAGCTAVVAPASQTLPLAL 190
 QY 188 VLAKAPAEAGLPAGVFNITGRSEIGDYIVHEEVNFINTGSTPVGQRIGLAG--NR 245
 191 KLAEHLAHEADIPRGVINITYGSAKALADYMEDEGRKXVSFTSTIGKELMASAQTMK 250
 DB 246 PIMLEIGKDGAGIVLADADLDNNAKOIVAGAYDSGORCTAIRVLVEEVADELAEKIS 305
 251 KVSLEIGGHAFFIVMNDADLDKAVEAVIGSKPRNAQOTCTCTRVPVQEEVYAPVEKFE 310
 QY 306 ENYAKLSVGDPF--DNATVTPVIDNSADFIESLVDAARQKAKELNEFRK----DGRLLT 360
 311 KAVGQKLKVDGFGDGTTPVPLIDENAVSKVQEHIBDAIQGQVLYGQGVAAELDGHFMQ 370
 DB 361 P---GLFDHTVTLDMKLAMEEPFGPIPIIRVDAEEVAVALANKSDPGLQSSVFTTRDFOKA 417
 371 PVIYGLANDTMLCMN---EETFGPVAPVAKFKTVEEIVIRAHNTPPGLAAVIFTKDISQA 427
 QY 418 FDIANKLEVGTVHINKTKRGPD--NPPLGLKSGAGVQGIIRYSIEAMTNVKSIVLDM 474
 428 FOISEALEYGIIGLNDGL---PSVAQAPFGGFGESGIGREGGHGIEEYLEIKYISLGL 483

RESULT 65
 DHAL_BACST
 ID DHAL_BACST STANDARD; PRT; 488 AA.
 AC P42329;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Aldehyde dehydrogenase, thermostable (EC 1.2.1.3).
 GN Name=alDH;
 OS Bacillus stearothermophilus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 OC NCBI_TaxID=1422;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=SI1;
 RA Imanaka T., Ohta T., Sakoda H., Widyasutji N., Matsunaka M.,
 RT "Cloning, nucleotide sequence, and efficient expression of the gene
 coding for thermostable aldehyde dehydrogenase from *Bacillus*

RT steaerothermophilus, and characterization of the enzyme.";
RL J. ferment. Bioeng. 76:161-167(1993).
CC -1- FUNCTION: Oxidizes several aliphatic aldehydes, particularly c6-
CC aliphatic aldehyde and hexanal, but do not oxidize benzaldehyde.
CC -1- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC -1- COFACTOR: Requires either NAD or NADP as a coenzyme.
CC -1- PATHWAY: Ethanol utilization; second step.
CC -1- SUBUNIT: Homotrimer.
CC -1- MISCELLANEOUS: The optimum temperature is around 55 and 60 degrees
CC Celsius.
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-slb.ch/announce/>
CC or send an email to license@1sb-slb.ch).
CC -----
CC EMBL: D13846; BAA02975.1; -
CC PIR: I39769; I39769.
CC HSSP: P51977; 1BXS.
CC InterPro: IPR002086; Aldehyde_dehydr.
CC Pfam: PF00171; Aldehyd; 1.
CC TIGRPFAMs: TIGR01804; BADH; 1.
CC TIGRPFAMs: TIGR01722; MMSDH; 1.
CC TIGRPFAMs: TIGR01780; SSADH; 1.
CC PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; 1.
CC NAD: Oxidoreductase.
CC NP BIND 233 NAD (ADP part) (By similarity).
CC FT ACT_SITE 255 255 By similarity.
CC FT ACT_SITE 289 289 By similarity.
CC SEQUENCE 488 AA; 52915 MW; 2982445196509DE CRC64;

Query Match 30.2%; Score 721.5; DB 1; Length 488;
Best Local Similarity 34.4%; Pred. No. 4.3e-35;
Matches 166; Conservative 104; Mismatches 192; Indels 21; Gaps 9;

QY 3 KEYQYNGCEKSSV-NQIE-ILSPIDSSL-GFVPMATREYVDHAKAGREALPMAAL 59
DB 8 KTYFYNGNWNVSSVNNVPSINPARNHDIYGVYRSTLEVDNEAVTANERQTSWMKR 67
QY 60 TYVERQVYHKADIIEERKKEIATVLAKEISKAVNASTVEVTRADLIRYAAEEGIRLS 119
DB 68 SGVEREEVLYKAAHILEQCLQDIATMTREMGKTLAEAKAETVRGVHILRYAGEGAR-- 125
QY 120 TSADEGKMDASTGHTLAVIRROVGIYLAIPYNYPNVLSGSKIAPALIGGVVFKFP 179
DB 126 ---KIGDVI PSSDSBELFTTRVPLGVGVISIPMNFPVALPIWKMAPALVGTIVLAKRA 182
QY 180 TGGSVGLVLAFAEAGLPAGVFNITTRGSEIGDYIVHEBEVNFINTGSTPVGORIG 239
DB 183 SETAVLAAYKVICFHEAGFPKGVNNVCGSGVAVCGIANNHPIDVFTFGSTVYKQVG 242
QY 240 KLAQMR--PIMLEIGSKDAGIVLADADLNNAKQIYAGAVDVSQGCRTAKRVLVBEVA 297
DB 243 RAAFERGAAYOLEMGSKNPVIYAKDADLDAVEGTISGGRSTGRCKTATSRPFIERYEV 302
QY 298 DELAEKISENVAKLSVGDPEFNAT-VTPYIDNSADFIESLVVDARQAKELNEFKR-- 354
DB 303 EEPFKAKLIRVYKQLKTGNGLDAETWGPCASBSQFHTVLSYIEKGSBKALKLYGNRCL 362
QY 365 -----DGRLLTPEGLPDHVTLDKMLAMEBPPGILPIIRYKDAEVAVALANKSDPGLQSSV 409
DB 363 EESLNGFVEVEPTIFEDVDLQMTIAREEIRFGPVLAIQVDSIEAIIKLANDREYGSASI 422
QY 410 FTTRDFKADIANKELEVGVTHINNKTRGRDNPFFGLK--GSGAIVGGRISYIEAMTV 467
DB 423 YTKNIGNALEFTKIDIEAGLIKVAETAGVEPQAPFGKQSSSHSREQG-QAAIEFTSI 481
QY 468 KSI 470

DB 482 KTV 484
|:|
RESULT 66
Q92DA3 PRELIMINARY; PRT; 488 AA.
ID Q92DA3;
AC Q92DA3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lin0913 protein.
GN OrderedLocustNames=lin0913;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_Taxid=1642;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser P., Frangeul L., Buchrieser C., Ruenik C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Darvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Eutlian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurupkat G.,
RA Madeno E., Maitouran A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tietrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species";
RL Science 294:849-852(2001).
DR EMBL: AL596166; CAC96145.1; -.
DR PIR: A11546; A11546.
DR HSSP: P05091; 1004.
DR ListerList; Lin0913; -.
DR GO: GO:0016491; P:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002086; Aldehyd_dehydrog.
DR Pfam: PF00171; Aldehyd; 1.
DR Complete Proteome.
SQ SEQUENCE 488 AA; 53351 MW; 3270FF2486CB08CE CRC64;

Query Match 30.1%; Score 720; DB 2; Length 488;
Best Local Similarity 33.6%; Pred. No. 5.3e-35;
Matches 160; Conservative 100; Mismatches 198; Indels 18; Gaps 6;

QY 8 YVNGCEKSSVNO--IEILSPIDSSLGFPVPMATREYVDHAKAGREALPMAALTVYERA 65
DB 16 FINGMKGDNKERTDYNPANGVIYAKIQAQPRETQALIKAKADAPDMAKMLARV 75
QY 66 QYIHRADIIERKKEIATVLAKEISKAVNASTVEVTRADLIRYAAEEGIRLSADEG 125
DB 76 KLIHNIADLMEBRADTLAKIMTLEQCKPLSKSGEVLTCGAENFRPAABEARLL-----YG 130
QY 126 GMDASTGHTKLAVIRROVGIYLAIPYNYPNVLSGSKIAPALIGGVVFKFPQSSVS 185
DB 131 ETTPANNHAF-IVKKQPIGVAAALTTPNPFQGMVTRKLAAPLATGNTVLPSSDTPLS 189
QY 186 GLVLAFAEAGLPAGVFNITTRGSEIGDYIVHEBEVNFINTGSTPVGORIGTAG-- 243
DB 190 ALAIFRIFEEAGLPKGVANNIYWGSSKEIGFTLTDSDDVRKLTFTOSTVVGOTLFGQSDT 249
QY 244 MRPIMLEIGKDGAGIVLADADLNNAKQIYAGAVDVSQGCRTAKRVLVBEVADELAEK 303
DB 250 LKSLIEIGGHPFIVFADANLEAAVDLITAKFRNNGQVCSPIRIVAKEIKEKFTKT 309
QY 304 ISENVAKLSVGDPEFNATVTPYIDNSADFIESLVVDARQAKELNEFKR-----DG 356
DB 310 IVAKEKIKVNGGLDGVNVGPIREDADIKIDKQIKNATDGAAYLVGTGERLTGSDYDKG 369
QY 367 RLITPEGLPDHVTLDKMLAMEBPPGILPIIRYKDAEVAVALANKSDPGLQSSVFTTRDFOK 416

Db 370 NFKPFTVLDNVTBMDIFYEETGFPVPLVFNEDDAIMKADSEGLASYITTDOLAR 429
QY 417 AFDIANKLEVGTVHINNKTGRGPDNFPFLGLKSGAGVGGIRYSIEMTNVKSIVL 472
Db 430 VEKVAALAEYGVGNANIEAISNPET-PFGVGKHSFGFRENHGHGMEEYIQVKFINL 484

RESULT 67

Q63GQ3 PRELIMINARY; PRT; 483 AA.
AC 063GQ3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DT Succlinate-semialdehyde dehydrogenase (NAD(P)+) (EC 1.2.1.16).
GN Name: gadi; ORFNames=BTZK0299;
OS Bacillus cereus ZK.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=288681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZK;
RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus cereus ZK";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000001; AAV19940.1; -.
KW Oxidoreductase.
SQ SEQUENCE 483 AA; 52243 MW; 30C61FA3F1420849 CRC64;

Query Match 30.0%; Score 718.5; DB 2; Length 483;
Best Local Similarity 35.8%; Pred. No. 6,5e-35;
Matches 170; Conservative 83; Mismatches 203; Indels 23; Gaps 8;

QY 8 YVNGEKSSVNOQIEILSPIDSSIGFVPANTREEDHAMKAGRALPAMALTYVERAQY 67
Db 16 YINGEWITTLQEOIEVNNPATKEIFATVPKGVTEAKQAVDAHAHAFKSWSKLTRAADRAK 75
QY 68 LHKRAADIIBDKKEIATVLAKEISKAYNASVTEVVRADLIRYAEERGLRSLTSADGCK 127
Db 76 LKRMFTLIDENKEEILAIIMTKEOCKPEPAALGEVNTANSFVEWYAEGRKV- - -YGEW 130
QY 128 MDASTGHLAVIRROPVGIYLAIAFYVYNLSGSKIAPIALIGNVVMFPTQGSVSGI 187
Db 131 IPASHNPKRILVWKOPGVAAITTPWNPFAAMITRKVAPALAGCTAAVVPASQTLTAL 190
QY 188 VLAKAPAEAGLPAGVNTITGRGSEIGDYIVEHEVNFINTGSTPVGORIGKLAG- -MR 245
Db 191 KLAEHLAHEADIPKGVINIVTGSAAKAIADTMEDGRVAKVSGFTSGTEIGKELMASAQTWK 250
QY 246 PIMLELGKQAGIVLADADIDNNAKQIVAGAYVYSGORCTAIRKVLVEEVADELAEKIS 305
Db 251 KVSLELGGAHPFTVMNDADIDKAVEATVIGSKFRNAGTCTCTNRVFOEVEVYAEVKEQK 310
QY 306 ENYAKLSVGDPE-DNATVTPVIDNSADFIESLVDARQKAEINLEFK- - -DGRILTL 360
Db 311 KAVGOLKVGDFGFDGTTVGPLIDENAVSKVQEHIEBILQKGVIVLVGGQVVALDGHFMG 370
QY 361 P- - -GLPDHVTLDMLKAMEEPGPILPIIRYKDAEBAVAIANKSPFGLOSSVFTDPQKA 417
Db 371 PVIIGLANDMTLMCMN- - -EETFGPVAVPAKFTVDEIERANHTPYGLAAVIFTKDISQA 427
QY 418 FDIANKLEVGTVHINNKTGRGPD- -NPPFLGLKSGAGVGGIRYSIEMTNVKSIVLDM 474
Db 428 FOISEALIEYIGLNDGL- - -PSVAQAPFGGFGYBSGIGREGGHPGIEVEYIEIKYISLGL 483

RESULT 68
Q6Y819 PRELIMINARY; PRT; 488 AA.
ID 06Y819
AC 06Y819;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE lmo0913 protein.
GN Ordered locus names=lmo0913;
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;

RX Glaeser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
RX Baquero F., Berche P., Bioecker H., Brandt P., Chakraborty T.,
RX Charbit A., Cheuonani F., Couve E., de Darvar A., Delhoux P.,
RX Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
RX Eutlier K.-D., Fehli W., Garcia-del Portillo F., Garrido P.,
RX Gautier L., Goebel M., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
RX Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
RX Madueno E., Maltournam A., Mata Vicente J., Ng E., Nedjati H.,
RX Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RX Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RX Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;
RT "Comparative genomics of Listeria species";
RL Science 294:849-852(2001).
DR EMBL; AL591977; CAC989991.1; -.
DR PIR; A11188; A11188.
DR HSSP; P05091; 1004.
DR List:List; lmo0913; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR Pfam; PF00171; Aldehdh; 1.
KW Complete proteome.
SQ SEQUENCE 488 AA; 53166 MW; 8EA0A920B1CE02A2 CRC64;

Query Match 30.0%; Score 718; DB 2; Length 488;
Best Local Similarity 33.8%; Pred. No. 7e-35;
Matches 161; Conservative 96; Mismatches 201; Indels 18; Gaps 6;

QY 8 YVNGEKSSVNO- -IEILSPIDSSIGFVPANTREEDHAMKAGRALPAMALTYVERA 65
Db 16 FINGKMTDQDNKETKDIYVNPANGDYAKIAQGPSETTKAIRAKADAFDMAKMELADYV 75
QY 66 QYLRKADIIBDKKEIATVLAKEISKAYNASVTEVVRADLIRYAEERGLRSLTSADGCK 125
Db 76 KLHKRIADIMBERKADTLAKIMTLEQCKPLKESKGEVLTVGVNFRFAEABARL- - -YG 130
QY 126 GKMDSTGHLAVIRROPVGIYLAIAFYVYNLSGSKIAPIALIGNVVMFPTQGSVSGS 185
Db 131 ETIPAPNNHAF-IVKKQPIGVAAITTPWNPFGAMVTRKLAIPALATGNTIVLRKSGDTPLS 189
QY 186 GLYLAKAPAEAGLPAGVNTITGRGSEIGDYIVEHEVNFINTGSTPVGORIGKLAG- - 243
Db 190 ALAIFEIFEAGLPKGVANITVWSSKEIGETLTSDDVAKLTFSTGKVGQTLFPKQADT 249
QY 244 MRPIMLELGKQAGIVLADADIDNNAKQIVAGAYVYSGORCTAIRKVLVEEVADELAEK 303
Db 250 LKKSILELGGAHPFTVFDANIDAVNDVAKKFRNNGVCSPPRIRFAKKEKEFTYA 309
QY 304 ISENVAKLSVGPDPNATVTPVIDNSADFIESLVDARQKAEINLEFK- - -DG 356
Db 310 LVAKVEQLKVGNGLDVNVGPIRPAIDIKOIKNAIEKAKAVLTGGGRITGSDYDYG 369
QY 357 RLTLPGFPHVTLDMLKAMEEPGPILPIIRYKDAEBAVAIANKSPFGLOSSVFTDPQKA 416
Db 370 NFKPFTVLDNVRKNDIFYEETGFPVPLVFNEDDAIMKADSEGLASYITTDOLAR 429
QY 417 AFDIANKLEVGTVHINNKTGRGPDNFPFLGLKSGAGVGGIRYSIEMTNVKSIVL 472
Db 430 VEKVAALAEYGVGNANIEAISNPET-PFGVGKHSFGFRENHGHGMEEYIQVKFINL 484

RESULT 69

```

OQ8X950
ID OQ8X950 PRELIMINARY; PRT; 482 AA.
AC OQ8X950; O7ABE9;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Succinate-semialdehyde dehydrogenase, NADP-dependent activity.
GN Name=gabd; OrderedlocusNames=EC63522, z3959;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxId=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 050952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakiwa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -i- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL; AB005495; AAG57768.1; -.
DR EMBL; AP002562; BAB36945.1; -.
DR PIR; B91069; B91069.
DR PIR; D85913; D85913.
DR HSSP; P20000; 1A68.
DR HSSP; Q28399; 1O9J.
DR GO; GO:0004028; F:aldehyde dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002086; Aldehyd_dehydrog.
DR InterPro: IPR010102; SSADH.
DR Pfam; PF00171; Aldedh; 1.
DR TIGRfams; TIGR01780; SSADH; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Oxidoreductase; Complete proteome.
SQ
SEQUENCE 482 AA; 51817 MW; 0CF3D4790A9EAD15 CRC64;

Query Match 30.0%; Score 717; DB 2; Length 482;
Best Local Similarity 33.9%; Pred. No. 8e-35;
Matches 162; Conservative 92; Mismatches 208; Indels 16; Gaps 7;
QY 6 QYVNGEMKSSVY--OIEILSPIDDSLGFVPMTREBYDHAKAGREALPMAALTYVE 63
DB 12 QALINEEMEDDANNGEVIDVTPNANGDKLSVPRMGADETPAALDAANRALFVFRALTAK 71
QY 64 RAQYLKAAADIIRDEKEIATVLAKEISKAYNASVTEVVRTADLIRYAAEGIRLSTAD 123
DB 72 RANILANWEMIMMEHODDLARLMTLLBOGKPLAEAKGEISYASFIEMFAEBGKRTYGDIT 131
QY 124 EGGKMDASTGHKLAIVIRQPVGIVLAIPYNYEVLNLSGSKIAPALIGGVNMFKPTQGS 183
DB 132 PGHQAD-----KRLIYIKQPIGVTAITPNNFPAAMITRKAGPALAAGCTMVLKPSQTP 186
QY 184 VSGVLTAFAFAEGLPAGVFNITITGSGSEIGDIYIEHEEVNFINFPGSTPVGORIGKLAG 243
DB 187 FSLALAEALIRAGIPAGVFNVTGSAGAVNGLTSPNPLVRKLSFTGSGTEIGRLMEQCA 246

```

```

QY 244 --MRPIMELGKGADGIADADLDNAKOIIVAGADYSGORCTAIKRLVVEEVADELA 301
DB 247 KDIKVYSLELGNNAFIVFDADLDKAVEGALASFRNAGGTCCANRLYVQDGYDIFA 306
QY 302 EKISNVAKLSVGDFPD-NATVTPYIDNSADFLIESLVADRAOKAGEL---NEFKRDGR 357
DB 307 EKLQAVSKLHIGDLDGVTIGPLIDEKAVAKBEHIALDLBKARVAVCGKAKHERGDN 366
QY 358 LITPLFPHVTLDMKLAMEEPRGPTLPIIRKDAEENAVAINKSPFGIQQSVFTFRPQKA 417
DB 367 FFOPTILVDVPAANAVSKEETFGPLAPLFRFDEADVIAQANDIEFGALAAFYADLSRV 426
QY 418 FDIANKLEGVTHINNKTG-RGPDNFPPLGKSGAGVQGIIRYSIEAMTNVKSIVLDM 474
DB 427 FRVGALRYGVIGIN--TGIISNEVAPRGGIKASGLGREGSKYGIEDYLEIKYMGIGL 482

RESULT 70
OQ9K01
ID OQ9K01 PRELIMINARY; PRT; 493 AA.
AC OQ9K01;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Probable aldehyde dehydrogenase.
GN OrderedlocusNames=Ts0809;
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
NCBI_TaxId=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001; DOI=10.1038/35035069;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Fritsman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513(2000).
CC -i- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL; AL445065; CAC11938.1; -.
DR HSSP; Q28399; 1O9J.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro: IPR02086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldedh; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ
SEQUENCE 493 AA; 54782 MW; 2F6A79DD96B8E7F5 CRC64;

Query Match 29.9%; Score 716.5; DB 2; Length 493;
Best Local Similarity 34.6%; Pred. No. 8.e-35;
Matches 167; Conservative 97; Mismatches 191; Indels 27; Gaps 9;
QY 8 YNNGEM--KSSVNOIEILSPIDDSLGFVPMTREBYDHAKAGREALPMAALTYVRA 65
DB 6 YIDGMVNSSSGKTYDKSPVYGVYIGRFEATRDVRAALDAADATAMANDLASSVERS 65
QY 66 QYLKRAADIIRDEKEIATVLAKEISKAYNASVTEVVRTADLIRYAAEGIRLSTADEG 125
DB 66 KLIYAKELIENKRAELNIEENKGVKKEKEVDVIDIQIYVAMARKLNGEYVEG 125
QY 126 GMDASTGHKLAIVIRQPVGIVLAIPYNYEVLNLSGSKIAPALIGGVNMFKP--PTQGS 183
DB 126 -----TSGHRKTFQYKVPYGVIVALTPNMFPAGVARKIAPALLTGNVTVLKPSDTPGS 180
QY 184 VSGVLTAFAFAEGLPAGVFNITITGSGSEIGDIYIEHEEVNFINFPGSTPVGORIGK--L 241
DB 181 AEWY--RKEVFAEPKGVLANFITGRSGSEIGDIYIEHKKVNLITMTGSTATGQRMQAS 238
QY 242 AGMRPIMELGKGADGIADADLDNAKOIIVAGADYSGORCTAIKRLVVEEVADELA 301

```

```

Db      239 ANNAKILIELGCGAPPMWKADMDNALKTLMAKWMNAGOSCIABRLVYHEDIYDTFM 298
      302 EKISENVAKLISVGDPPDNATVTPEVIDNSADFIESLVDAKQKAK-----ELNEFK 353
Qy      299 SRFEVLSRKALGDP--KNADMGPLINKGALQATSEIVEEAKESGAKILFGSGQPSLSGTY 357
Db      354 RQGRLLTPGLFDHVTLDMLAMEEPFGPIPIIRVYDAEBAVALANKSPFGLOSSVFTPD 413
      358 RUGYFPLFTIGADQSKIFQSEIFAPYIGARKISSVEEMYLADLSKGLASLYLPTXD 417
Qy      414 FQKAPFIAKLEVGTHINNKTRGPDNFP--FLGKSGAGVQGI RYSEAMTNVKSIV 471
Db      418 PNIFPASERIRFELGYVMP---GPEASGQHTGFRMTGQAGSGSKYGISYLKKNIV 474
Qy      472 LD 473
      475 VD 476

```

RESULT 71

```

ID      081DV8      PRELIMINARY;      PRT;      474 AA.
AC      081DV8;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Succinate-semialdehyde dehydrogenase (EC 1.2.1.16).
GN      OrderedLocNames=BC2241;
OS      Bacillus cereus (strain ATCC 14579 / DSM 31).
OC      Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
NCBI_TaxID=226900;
RN      NCB1_TaxID=226900;
RX      SEQUENCE FROM N.A.
RP      MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA      Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA      Kapral V., Bhattacharya A., Reznik G., Mikhailova N., Japids A.,
RA      Chu L., Mazur M., Goldsman E., Larsen N., D'Souza M., Walunas T.,
RA      Grecklin Y., Pusch G., Haselkorn R., Fongstein M., Ehrlich S.D.,
RA      Overbeek R., Kyrides N.C.;
RT      "Genome sequence of Bacillus cereus and comparative analysis with
RT      Bacillus anthracis.";
RL      Nature 423:87-91(2003).
DR      EMBL; AE017005; AAP09206.1; -.
DR      HSSP; Q28399; 109J.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0009013; F:succinate-semialdehyde dehydrogenase [NMD (P. . .); IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR002086; Aldehyd_dehydrog.
DR      Pfam; PF00171; Aldehyd. 1.
KW      Complete proteome; Oxidoreductase.
SQ      SEQUENCE 474 AA; 51868 MW; 598AE1E3B1D5C8 CRC64;

```

Query Match 29.9%; Score 716; DB 2; Length 474;

Best Local Similarity 35.1%; Pred. No. 9e-35;

Matches 165; Conservative 98; Mismatches 199; Indels 8; Gaps 5;

```

Qy      8 YVNGEKSSVNOEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMAALTYVERAY 67
      6 YVNGEKSSVNTTYPLAIFYSEETLAIAGTSEEDVEKAVTAANKAMTKANKKSAVRAPI 65
Db      68 LHKAAIIRDRKEEIAVLAKESIKAYNASVTEVVRADLIRYAAEGIRLSTSADEGK 127
      66 LEKVAQMDRERREFEIILAKKAPRPARGEVDRTVQYKFAAEBAKRI---YESTLP 122
Qy      128 MDASTGK--KTLAIVRQPGVIVLAIPYNPVNLSSGSKAPALIGNVVMFKPPTQGSYS 185
Db      123 LBAAPAGDGIATITRKIPGIVICATIPPNPLLVAKHVGPAIAAGTVAALKRADDTPIS 182
Qy      186 GLVLAFAFAEAGLPAGVFNTITRGSEIGDYIVEHEEVNFINTGSPVQORIGKLAKMR 245
Db      183 SYALVELFERAGLPNGALNTIISGPGTVGSAIVKNDVVASITTTGSPKVGIGIKQKAGLK 242

```

```

Qy      246 PIMLEIGKADGIVLADADL--DNAKOIVAGAYDYGSGRCAITKRVLVVEEVADELEKTI 304
      243 RVTLEIGSNAVITIEDVEHLEITERVKMGAFVANNQVCISQVRFVHETIKIHELSTG 302
Qy      305 SENVAKLISVGD--PNNATVPYIDNSADFIESLVDAKQKAKEL--NEFKDGRLLTPG 362
Db      303 KRAMSVVVGDPISLEETDVSAIISKDVERIDMWQEAIKEGATVLCGKKQDARIFEBPT 362
Qy      363 LFDHTLMDKLAMEEPFGPIPIIRYKDAEEVAVALANKSPFGLOSSVFTPDQKAFDIAN 422
      363 VLTNPDPHVSVOCEVFGPLMTVNTFKERDEALIEQVNNSRYGLOAGVFNTNLFKMRARD 422
Qy      423 KLEVGTHINNKTRGPDNFPFLGLKSGAGVQGI RYSEAMTNVKSIVL 472
Db      423 ELEVGVMINDIPTFRVDHMPYGVKESGTGREGIKYALTEEMTKLVCI 472

```

RESULT 72

```

ID      09HQ22      PRELIMINARY;      PRT;      480 AA.
AC      09HQ22;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Glyceraldehyde-3-phosphate dehydrogenase.
GN      NameGap; OrderedLocNames=VNG0937G;
OS      Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC      Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC      Halobacteriaceae; Halobacterium.
NCBI_TaxID=64091;
RN      NCB1_TaxID=64091;
RX      SEQUENCE FROM N.A.
RP      MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA      Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA      Shukla H.D., Lasky S.R., Balliga N.S., Thorsen V., Shrogon J.,
RA      Swartzell S., Weir D., Hall J., Dahl T.A., Meli L., Goo Y.A.,
RA      Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA      Maddocke D.G., Jablonski P.B., Krebs M.P., Angevine C.M., Dale H.,
RA      Isebnarger T.A., Beck R.F., Pohlischroder M., Spindlich J.L., Jung K.-H.,
RA      Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA      Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT      "Genome sequence of Halobacterium species NRC-1";
RL      Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR      EMBL; AE005031; AAG19367.1; -.
DR      PIR; C84250; C84250.
DR      HSSP; O57693; 1KY8.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR002086; Aldehyd_dehydrog.
DR      Pfam; PF00171; Aldehyd. 1.
DR      PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN 1.
DR      PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN 1.
KW      Complete proteome.
SQ      SEQUENCE 480 AA; 50483 MW; FE86939702082CD2 CRC64;

```

Query Match 29.8%; Score 713.5; DB 2; Length 480;

Best Local Similarity 35.0%; Pred. No. 1.3e-34;

Matches 165; Conservative 86; Mismatches 210; Indels 11; Gaps 6;

```

Qy      8 YVNGEKSSVNOEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMAALTYVERAY 67
      13 FYAGEVNSDGEPLPVSDLDGGTFASVAADTSDAERLAAATGVHADLRETTVERVEM 72
Db      68 LHKAAIIRDRKEEIAVLAKESIKAYNASVTEVVRADLIRYAAEGIRLSTSADEGK 127
      73 LESIADGIRRRDELAELAVIRBAKPRISSARBEVGSAAERFRAVGE--LRHITGEVRTG- 130
Qy      128 MDASTGK--LAVIRQPGVIVLAIPYNPVNLSSGSKAPALIGNVVMFKPPTQGSYS 185
Db      131 --TTAGHEHMOAIVKHEPMTVLCITPPYPLSTALQVAPAIAGNAVIVPASTPTIS 188
Qy      186 GLVLAFAFAEAGLPAGVFNTITRGSEIGDYIVEHEEVNFINTGSPVQORIGKLAKMR 245
      183 SYALVELFERAGLPNGALNTIISGPGTVGSAIVKNDVVASITTTGSPKVGIGIKQKAGLK 242

```

```
Db 189 GAILADIADAGLPGAGVNFVESSVIGDPLASDARVDIAIMTSSSGAGEHVARQGIT 248
QY 246 PIMLELGGAGAGIYVADADLDNAKQIVAGAYDSGORCTAIKRVLVVEEVADELAEKIS 305
Db 249 RLHMLGGNAAPVIEVEDADLDNAADATRGSLKIVGORCSAVSVLAHESVDELVSRI 308
QY 306 ENVAKLVSQDPPD-NATYTPVIDNSADFIESLVVDARQKAKELNEFKR---DG-RLLT 360
Db 309 DAMAEMSIGDLFDITTLGFLVSADQADWVAELVDAVDRGATVVRGGERHVEDGVHYE 368
QY 361 PELFPHVYLDKMLAMEEPFGPIPIIRVQDAEAAVAIAKSPFGLOSVFTFDQKAFPI 420
Db 369 PFLADVPDARIIVDEEGPVCAYTVTVDEDDAVRTANGSELALDAAVFTADHDRAMRV 428
QY 421 ANKLEVGTVHINNKTRGPDNFPFLGLKSGAGVQGIIRYSIEAMTNVKSIVL 472
Db 429 AERVAVAGVRINGAPSHGIDVFPFGNDASGIGREGDSTHEFVREKSIIL 480

RESULT 73
057693 PRELIMINARY; PRT; 501 AA.
AC 057693;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (NADP+) (EC 1.2.1.9).
GN Name=GAPN;
OS Thermoproteus tenax.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Thermoproteus.
OX NCBI_TaxID=2271;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KRA 1;
RA Brunner N.A.;
RT Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; Y10625; CAA71651.1; -.
DR PIR; T44939; T44939.
DR PDB; 1KX8; X-ray; A=1-501.
DR PDB; 1UXN; X-ray; A=-.
DR PDB; 1UXP; X-ray; A=-.
DR PDB; 1UXO; X-ray; A=-.
DR PDB; 1UXR; X-ray; A=-.
DR PDB; 1UXT; X-ray; A=-.
DR PDB; 1UXV; X-ray; A=-.
DR PDB; 1UXW; X-ray; A=-.
DR GO; GO:000886; F:glyceraldehyde-3-phosphate dehydrogenase (N. .; IEA.
DR GO; GO:001649; F:oxidoreductase activity; IEA.
DR InterPro; IPR002086; Aldehyde_dehydrog.
DR Pfam; PF00171; Aldehyd; 1.
DR KEGG; OXidoreductase.
SQ SEQUENCE 501 AA; 54090 MW; 22F1C0084282EB1B CRC64;

Query Match 29.8%; Score 713; DB 2; Length 501;
Best Local Similarity 37.0%; Pred. No. 1.4e-34;
Matches 180; Conservative 83; Mismatches 196; Indels 28; Gaps 9;

QY 5 YONYNGEMKSSVNVNIEILSPIDSSLGFPVAMTRREVDHAM-----KAGREALPMAA-- 58
Db 19 YPSYIAGEGGSGGQIEVKSPLDLTIAVISPSSREVRRTLDVLFKGR-----WSARD 73
QY 59 LTVVERAOYLHKRAADIIEERDKKEIATVLAKEISKAVNASVTEVVTADLIRYAAEGRI 118
```

```
Db 74 MPTGERLAVLNRKADIIERNLDVFAEVLWNAGKPKSAVGEVRAADVRL-LAEFLDKK 132
QY 119 STSADGGKMDASIGCHKLAIVRRQPVGIYLAIPNTYVNISSKIALPAGNVWMEKP 178
Db 133 IGGDYIPDMDWTYDITLETBGLVRREPLGVAAITPNNYPLFAVVKITYSFTYGNAAVVKP 192
QY 179 PTQGSGLVTLAKAPAEAGLPAGVFNITTGSGSEIGDYIVHEEVNFINTGSPFVGORI 238
Db 193 SISDPLPAAAVKALLDGFPDPAIALINLPQKE-AEKIVADDRVAVSTFGSTVEGRV 251
QY 239 GKLAMRPIMLELGGAGIYVADADLDNAKQIVAGAYDSGORCTAIKRVLVVEEVA 298
Db 252 VKVGSKQYVNMELGGDPAIVLEDDDLADKIARGIYSYAGQCDATKVLAEKRPYVG 311
QY 299 ELAEKISNVAKLSVQDPPD-NATYTPVIDNSADFIESLVVDARQKAKELNEFKRGR 357
Db 312 KLVEEVAARLSSLRVDRDPTVDVGPLISSAVDENMAALIDAVEKGRVLT---AGGR 367
QY 358 LITPLGPDHVTL-----DMKLAMEEPFGPIPIIRVQDAEAAVAIAKSPFGLO 408
Db 368 RLGP-TYIQPLIVERPADRVDMVLYKREVPAPVAVSEVVDLDQAITLANGRPYGLDAA 426
QY 409 VTRDFQKAFDIAKLEVGTVHINNKTRGPDNFPFLGLKSGAGVQGIIRYSIEAMTNV 468
Db 427 VFGRDVVKIRAVRLVEGAIYIDMPRHGIGYYPFGGRKKSQVFRREGIYGAVEAVTAYK 486
QY 469 STVDMK 475
Db 487 TIVFNK 493
```

```
RESULT 74
08EMH9 PRELIMINARY; PRT; 472 AA.
ID 08EMH9
AC 08EMH9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Succinate-semialdehyde dehydrogenase (EC 1.2.1.16).
GN OrderedLocusNames=O82863;
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments."
RL Nucleic Acids Res. 30:3927-3935 (2002).
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL; AP004602; BAC14819.1; -.
DR HSSP; P51977; 1BXS.
DR GO; GO:0004028; F:aldehyde dehydrogenase activity; IEA.
DR GO; GO:001649; F:oxidoreductase activity; IEA.
DR GO; GO:0009013; F:succinate-semialdehyde dehydrogenase [NAD(P). .; IEA.
DR GO; GO:0008152; F:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydrog.
DR InterPro; IPR010102; SSADH.
DR Pfam; PF00171; Aldehyd; 1.
DR TIGRFAMs; TIGR01780; SSADH; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR Complete proteome; Oxidoreductase.
SQ SEQUENCE 472 AA; 51376 MW; 7DCBDB7E0F0D552 CRC64;

Query Match 29.7%; Score 711.5; DB 2; Length 472;
Best Local Similarity 34.8%; Pred. No. 1.7e-34;
Matches 167; Conservative 91; Mismatches 201; Indels 21; Gaps 7;

QY 3 KEYQNVNNGEM-KSSVNVNIEILSPIDSSLGFPVAMTRREVDHAMKAGREALPMAALTV 61
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2005, 10:04:24 ; Search time 167 Seconds
(without alignments)
1100.067 Million cell updates/sec

Title: US-09-868-195-12

Perfect score: 2393
Sequence: 1 LTKKEYQNYNGEMKSVNOI.....GIRYSLEANTYKSYLDWK 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 80 summaries

Database : A_Geneseq.16Dec04:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2382	99.5	475	5	ABP27643 Streptococcus
2	2380	99.5	475	3	AAV96743 Streptococcus
3	2373	99.2	475	5	ABP29820 Streptococcus
4	1812	75.7	474	6	ABU01538 S. pneumo
5	1812	75.7	474	6	ABP81589 Streptococcus
6	1808.5	75.6	474	8	ADK47293 Streptococcus
7	1802	75.3	474	6	ABU46050 Protein e
8	1700	71.0	475	6	ABU44392 Protein e
9	1699	71.0	439	4	AU037723 Streptococcus
10	1692	70.7	436	8	ADR95169 Novel S.
11	1641	68.6	475	5	ABP27644 Streptococcus
12	1641	68.6	475	8	ADR83939 S. pyogen
13	1639	68.5	475	6	ABU46701 Protein e
14	1397	58.4	486	6	ABU25619 Protein e
15	1388.5	58.0	486	6	ABU24815 Protein e
16	1383	57.8	482	6	ABU24149 Protein e
17	1335	55.8	472	8	ADS27503 Bacterial
18	1312	54.8	473	8	ADS27631 Bacterial
19	1244	52.0	481	8	ADS28268 Bacterial
20	1172.5	49.0	498	3	AAAG0912 Zea mays
21	1169.5	48.9	496	7	ABM74113 DNA clone
22	1168.5	48.8	496	3	AAAG32607 Arabidops
23	1168.5	48.8	496	3	AAAG9550 Arabidops
24	1168.5	48.8	496	8	ADN73355 Thale cre
25	1164.5	48.7	496	3	AAAG05753 Arabidops

26	1128.5	47.2	461	3	AAAG32608 Arabidops
27	1128.5	47.2	461	3	AAAG9551 Arabidops
28	1103.5	46.1	438	3	AAAG9552 Arabidops
29	1103.5	46.1	438	3	AAAG32609 Arabidops
30	1102.5	45.9	438	3	AAAG40913 Zea mays
31	1099.5	45.9	438	3	AAAG05754 Arabidops
32	1036.5	43.3	402	3	AAAG05755 Arabidops
33	868.5	36.3	475	6	ABU48897 Protein e
34	786.5	32.9	492	6	AAE37093 Human 913
35	784.5	32.8	293	3	AAAG40914 Zea mays
36	777	32.5	493	7	AAE39889 Human ald
37	777	32.5	493	8	ADQ48417 Aldehyde
38	763	31.9	488	8	ADS44637 Bacterial
39	759.5	31.7	463	8	ADN18420 Bacterial
40	741.5	31.0	455	8	ADN18535 Bacterial
41	741.5	31.0	484	4	AAE79356 Coryneb
42	741.5	31.0	484	4	AAAG92676 C glutam
43	737	30.8	476	8	ADN17364 Bacterial
44	735	30.7	482	8	ADS28216 Bacterial
45	734	30.7	471	7	ADM26502 Hyperther
46	733.5	30.7	526	7	ABO78556 Pseudom
47	728	30.4	470	8	ADS27421 Bacterial
48	726	30.3	482	4	AAU34671 E. coli c
49	726	30.3	482	6	ABU28726 Protein e
50	726	30.3	482	6	ADS45081 Bacterial
51	722	30.2	511	8	ADN46591 Bacterial
52	721.5	30.2	483	6	ABU19008 Protein e
53	721.5	30.2	488	2	AAE55135 Heat res
54	718	30.0	488	5	ABBA7807 Listeria
55	718	30.0	488	6	ABU32914 Protein e
56	716.5	29.9	493	8	ADS29166 Bacterial
57	713.5	29.8	480	8	ADS29560 Bacterial
58	711	29.7	482	6	ABU31620 Protein e
59	711	29.7	487	7	ABO66116 Klebsiell
60	710.5	29.7	489	8	ADS29678 Bacterial
61	705.5	29.5	473	8	ADS24538 Bacterial
62	705.5	29.5	491	5	ABBO8382 PCuC amin
63	705.5	29.5	491	6	ABU92062 Pseudom
64	704	29.4	480	6	ABU40285 Protein e
65	697.5	29.1	468	8	ADS28136 Bacterial
66	694	29.0	473	8	ADS29963 Bacterial
67	694	29.0	477	8	ADN26088 Bacterial
68	693	29.0	480	6	ABU41623 Protein e
69	693	29.0	480	6	ABU42152 Protein e
70	692	28.9	482	6	ABU47293 Protein e
71	691	28.9	478	8	ADS24924 Bacterial
72	690	28.8	489	6	ABU23300 Protein e
73	687	28.7	480	8	ADS22151 Bacterial
74	686	28.7	480	8	ADS25659 Bacterial
75	682.5	28.5	458	8	ADN24384 Bacterial
76	682.5	28.5	458	8	ADN21629 Bacterial
77	682.5	28.5	485	8	ADS28246 Bacterial
78	682	28.5	482	4	AAU38454 Salmonell
79	682	28.5	482	6	ABU47996 Protein e
80	680.5	28.4	460	8	ADS27648 Bacterial

ALIGNMENTS

RESULT 1
ID ABP27643 standard, protein, 475 AA.
AC ABP27643;
XX 02-JUL-2002 (first entry)
XX Streptococcus polypeptide SEQ ID NO 4462.
DE Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
XX Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
KW Group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

```

XX OS Streptococcus agalactiae.
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB004789.
XX PR 27-OCT-2000; 2000GB-00026333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Maignant V, Margarit Y Rost, Grandi G, Fraser C;
XX PI Tetteiln H;
XX DR WPI; 2002-352536/38.
XX DR N-PSDB; ABN68274.
XX PS
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for
XX PT detecting a compound that binds to the protein.
XX PS
XX PS Claim 1; Page 3607; 4525pp; English.
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX CC the specification. The proteins have antibacterial and antiinflammatory
XX CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
XX CC antibodies that bind (I) are used in the manufacture of medicaments for
XX CC the treatment or prevention of infection or disease caused by
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a
XX CC biological sample. (I) is used to determine whether a compound binds to
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX CC used as a vaccine or diagnostic composition. The disease caused by
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic
XX CC acid encoding (I) may be used to recombinantly produce (I) and may be
XX CC used in gene therapy. Antibodies to (I) are used for affinity
XX CC chromatography, immunoassay, and distinguishing/identifying
XX CC Streptococcus proteins
XX SQ
XX Sequence 475 AA;
XX
XX Query Match 99.5%; Score 2382; DB 5; Length 475;
XX Best Local Similarity 99.2%; Pred. No. 4, 5e-190;
XX Matches 471; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 LTRKYQNTYNGEMKSSVNOIEILSPIDDSLGFPVPMTEBEVDHAKAGREALPMAALT 60
QY :|||||
DB 1 MKKEYQNTYNGEMKSSVNOIEILSPIDDSLGFPVPMTEBEVDHAKAGREALPMAALT 60
QY 61 VVERAQLHKAADIIRDKBEIATVLAKEISKAYNSVTEVTRADLIRYAAEEGRILT 120
QY :|||||
DB 61 VVERAQLHKAADIIRDKBEIATVLAKEISKAYNSVTEVTRADLIRYAAEEGRILT 120
QY 121 SADEGGKMDASTGKHLAVIRQPVGIVLAIPAPYEPVNLGSKTIAPALIGANVMKPEPT 180
QY :|||||
DB 121 SADEGGKMDASTGKHLAVIRQPVGIVLAIPAPYEPVNLGSKTIAPALIGANVMKPEPT 180
QY 181 QGSVSGLVAKAPAEAGLPAGVFNITTGSGSEIGDIYVHBEVNFNFTGSPVQGRICK 240
QY :|||||
DB 181 QGSVSGLVAKAPAEAGLPAGVFNITTGSGSEIGDIYVHBEVNFNFTGSPVQGRICK 240
QY 241 LAGMPRIIMEIGSKAGIYLAADLNLNAKQIYAGAYDVGSGRCIAIKVLAVEEVADEL 300
QY :|||||
DB 241 LAGMPRIIMEIGSKAGIYLAADLNLNAKQIYAGAYDVGSGRCIAIKVLAVEEVADEL 300
QY 301 AEKISENVAKLISVGPFDNATVTPVIDNSADFIESLVVDARQGAKEINFEKRDRLIT 360

```

```

DB 301 AEKISENVAKLISVGPFDNATVTPVIDNSADFIESLVVDARQGAKEINFEKRDRLIT 360
QY 361 PGLFDHNTLDMKLAEEPPFGPLPIIRYKDAEEVAALNKSDFGQSSVFTRDPOKADI 420
DB 361 PGLFDHNTLDMKLAEEPPFGPLPIIRYKDAEEVAALNKSDFGQSSVFTRDPOKADI 420
QY 421 ANKLEVGTVHINNKTRGPDNPPFLGLKSGAGVQIGIRYSIEAMTNVKSIVLDMK 475
DB 421 ANKLEVGTVHINNKTRGPDNPPFLGLKSGAGVQIGIRYSIEAMTNVKSIVLDMK 475
XX
XX RESULT 2
XX AA96743
XX ID AA96743 standard; protein; 475 AA.
XX AC
XX AC AA96743;
XX DT 26-SEP-2000 (first entry)
XX DE
XX DE Streptococcus NADP-dependent glyceraldehyde-3-phosphate dehydrogenase.
XX OS
XX OS Streptococcus agalactiae.
XX FH Key Location/Qualifiers
XX FT Misc-difference 297
XX FT /note="Encoded by GCW"
XX FT Misc-difference 439
XX FT /note="Encoded by CCM"
XX PN WO200037490-A2.
XX PD 29-JUN-2000.
XX PF 22-DEC-1999; 99MO-GB004376.
XX PR 22-DEC-1998; 99GB-00028346.
XX PR 20-JAN-1999; 99GB-00001233.
XX PR 12-JAN-1999; 99GB-00001234.
XX PR 12-APR-1999; 99GB-00008321.
XX PR 24-MAY-1999; 99GB-00012036.
XX PR 23-SEP-1999; 99GB-00022596.
XX PA (MICR-) MICROSCIENCE LTD.
XX PI Hughes MJG, Santangelo JD, Lane JD, Feldman R, Moore JC;
XX PI Everest P, Dobson RJ, Henwood CJ, Dougan G, Wilson RJ;
XX DR N-PSDB; AA51293.
XX DR WPI; 2000-442636/38.
XX PS
XX PS Peptide encoded by operon including genes identified and obtained from
XX PT group B Streptococcus, for treating streptococcal infections.
XX PS Example 5; Page 31-32; 32pp; English.
XX
XX The S. agalactiae MS14 gene encodes a protein which has homology with a
XX CC non-phosphorylating, NADP-dependent glyceraldehyde-3-phosphate
XX CC dehydrogenase (NPGAP-3-DH). Peptides derived from group B Streptococcus
XX CC (Streptococcus agalactiae) proteins encoded by genes MS4, MS10, MS11,
XX CC MS14 and/or MS16 are claimed. The peptides are useful for screening
XX CC potential drugs, or for the detection of virulence, and for the
XX CC manufacture of a medicament for use in the treatment or prevention of
XX CC infections such as focal infection (including osteomyelitis, septic
XX CC arthritis, abscesses and endophthalmitis) and urinary tract infections
XX CC caused by group B Streptococcus. The peptides and vaccines comprising the
XX CC peptides are useful in treatment of chronic mastitis, especially in cows
XX CC (i.e. for veterinary purposes)

```


SQ Sequence 475 AA;

Query Match	99.5%	Score 2380;	DB 3;	Length 475;
Best Local Similarity	99.6%	Pred. No. 6.6e-190;		
Matches 473; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy	1	LTELEYONYVNGEKSXSVNOIELLSPDDSGLFVAMTREEDYDAMKGRALPAMALTT	60
Db	1	LTELEYONYVNGEKSXSVNOIELLSPDDSGLFVAMTREEDYDAMKGRALPAMALTT	60
Qy	61	VYERAOYLKHAADIIERDKEEIATYLAKEISKANVASVEYVRADLIRYAAEEGIRLST	120
Db	61	VYERAOYLKHAADIIERDKEEIATYLAKEISKANVASVEYVRADLIRYAAEEGIRLST	120
Qy	121	SAEGBGKMASTGHKLAVTRRQPVGIYALAPYNNPVNLLSSSKTAPALIGNVMMFEPPT	180
Db	121	SAEGBGKMASTGHKLAVTRRQPVGIYALAPYNNPVNLLSSSKTAPALIGNVMMFEPPT	180
Qy	181	QGSVSGVLVAKAPAEAGLPAGVFNITRGSREIGDYIYEHEEYVNFIFNTGSPVQGRIGK	240
Db	181	QGSVSGVLVAKAPAEAGLPAGVFNITRGSREIGDYIYEHEEYVNFIFNTGSPVQGRIGK	240
Qy	241	LAGMRPIMTELGGKQAGIVLADADIDNAAKOIVAGADYDSCQRCTAIRVLVEEVADL	300
Db	241	LAGMRPIMTELGGKQAGIVLADADIDNAAKOIVAGADYDSCQRCTAIRVLVEEVADL	300
Qy	301	AEKISENVAKLSVGDPFDNATVTPVIDNSADPIESLTVVDRQKGAELNEFFKRDGILLT	360
Db	301	AEKISENVAKLSVGDPFDNATVTPVIDNSADPIESLTVVDRQKGAELNEFFKRDGILLT	360
Qy	361	PGLFDEHVTLDMKLAWEPPGPILPIIRIVYDAEBAVAIANKSDPGLQSSVFTRDFOKAFDI	420
Db	361	PGLFDEHVTLDMKLAWEPPGPILPIIRIVYDAEBAVAIANKSDPGLQSSVFTRDFOKAFDI	420
Qy	421	ANGLAEVGYTHINNKTRGPDNPPFGLKGSAGVCGIRISYEAMTNVYSIYLDNKK	475
Db	421	ANGLAEVGYTHINNKTRGPDNPPFGLKGSAGVCGIRISYEAMTNVYSIYLDNKK	475

XX	RESULT 3
XX	ABP29820
XX	ID ABP29820 standard; protein; 475 AA.
XX	AC
XX	ABP29820;
XX	02-JUL-2002 (first entry)
XX	DT
XX	DE Streptococcus pyopeptide SEQ ID NO 8816.
XX	DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX	KM group A streptococcus; Streptococcus pyogenes; antibacterial;
XX	KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX	OS Streptococcus agalactiae.
XX	PN WO200234771-A2.
XX	PD 02-MAY-2002.
XX	PF 29-OCT-2001; 2001WO-GB004789.
XX	PF 27-OCT-2000; 2000GB-0002633.
XX	PR 24-NOV-2000; 2000GB-00028727.
XX	PR 07-MAR-2001; 2001GB-00005640.
XX	PA (CHIR-) CHIRON SPA.
XX	PA (GENO-) INST GENOMIC RES.
XX	PI Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C,
XX	PI Tettelin H;
XX	DR WPI; 2002-352536/38.
XX	DR N-PSDB; ABN70451.

XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
PS Claim 1, Page 3999; 4525p; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/SBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
CC the specification. The proteins have antibacterial and anti-inflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN7156 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus (I) can be prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may
CC be used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins

SQ Sequence 475 AA;

Query Match	99.2%	Score 2373;	DB 5;	Length 475;
Best Local Similarity	98.9%	Pred. No. 2.5e-189;		
Matches 470;	Conservative 2;	Mismatches 3;	Indels 0;	Gaps 0;

Qy		1	LKEVONYNANGSMKSNVOJELISPIIDSSIGFPAWMPAMTREEDVHAMRAGEALPAMALIT	60
Dz		1	LKEVONYNANGSMKSNVOJELISPIDDSIGFPAWMPAMTREEDVHAMRAGEALPAMALIT	60
Qy		61	VYERAOYLHKAAADIIERDKEEIATVLAKESIKAVNASVTEVVRADLIIRYAABEGRIJLST	120
Dz		61	VYERAOYLHKAAADIIERDKEEIATVLAKESIKAVNASVTEVVRADLIIRYAABEGRIJLST	120
Qy		121	SADDEGKDNDASGHGLAIYIRROPGIVIALAPNYVPNLSSGTAPLIGANVMFPPT	180
Dz		121	SADDEGKDNDASGHGLAIYIRROPGIVIALAPNYVPNLSSGTAPLIGANVMFPPT	180
Qy		181	QGSVSGLYLAKAFAEAGLPACVFNTITGRGEIEDYIVEHEBNVNFINFSTSTPVQRIGK	240
Dz		181	QGSVSGLYLAKAFAGLPACVFNTITGRGEIEDYIVEHEBNVNFINFSTSTPVQRIGK	240
Qy		241	LGMRPIMLELGOGDAGIVLADADIDNNAKOIVAGAYDISGORCTAIRLVVEEVADEL	300
Dz		241	LGMRPIMLELGOGDAGIVLADADIDNNAKOIVAGAYDISGORCTAIRLVVEEVADEL	300
Qy		301	AARKISENNVAKLSVGPDPNATVTTPVIDNSADPIESI.VVAROKGAELNEFPKRDRILIT	360
Dz		301	AARKISENNVAKLSVGAPPDNATVTTPVIDNSADPIESI.VVAROKGAELNEFPKRDRILIT	360
Qy		361	PGLFPHVTLTDMKLAEBEPFGPILPIIRYKAOEBAVALANKSDPGLQSVFTRDFOKAFDI	420
Dz		361	PGLFPHVTLTDMKLAEBEPFGPILPIIRYKAOEBAVALANKDFGLQSVFTRDFOKAFDI	420
Qy		421	ANKLEVGVTHINNKTKRGPDNPFPPLGLKGSGAQOGRYSIEAMTNVKSIVFLDMK	475
Dz		421	ANKLEVGVTHINNKTKRGPDNPFPPLGLKGSGAQOGRYSIEAMTNVKSIVFLDMK	475

RESULT	4
ABU01538	
ID	ABU01538 standard; protein; 474 AA
XX	
AC	ABU01538;
XC	
DT	23-OCT-2003 (revised)
XX	
DT	11-FEB-2003 (first entry)
XX	

CC detecting S. pneumoniae in a biological sample or diagnosing S.
CC pneumoniae infection in a subject. The polynucleotides have antibacterial
CC activity and are useful in gene therapy

XX Sequence 474 AA;

Query Match 75.7%; Score 1812; DB 6; Length 474;
Best Local Similarity 74.1%; Pred. No. 1,8e-142;
Matches 349; Conservative 57; Mismatches 65; Indels 0; Gaps 0;

YY 5 YONVNGEKKSVNOIELISPIDSSLGFPVPMTEEVDMKAGREAPAMALTYER 64
DB 4 YONVNGEKKSVNOIELISPIDSSLGFPVPMTEEVDMKAGREAPAMALTYER 63
YY 65 AOVTHKADIIERDKERIATVLAKEISKAVNASVTEVTRADLIRAAEGIRLSTADE 124
DB 64 AAVLHTAALIERDKERIATVLAKEISKAVNASVTEVTRADLIRAAEGIRLSTADE 123
YY 125 GGGMDASTGHKLAVIRROPVGIATAPVNPVNLSSKIAAPALIGNVVMFKPPTQGSV 184
DB 124 GGGFEATSKNKLAVRRBPVGIATAPVNPVNLSSKIAAPALIGNVVMFKPPTQGSV 183
YY 185 SGLVAKAFAPAGAGCPAGVFNITGRSGEIGDYIVHEBEVNFINTGSTPVGQRIGLAGM 244
DB 184 SGLVAKAFAPAGAGCPAGVFNITGRSGEIGDYIVHEBEVNFINTGSTPVGQRIGLAGM 243
YY 245 RPIMLELGGKAGIATADADLNAKQIVAGAYDYGQRCTAIKRVLVVEVADLAEKI 304
DB 244 RPIMLELGGKAGIATADADLNAKQIVAGAYDYGQRCTAIKRVLVVEVADLAEKI 303
YY 305 SENVAKLVSQDPEDNATVPVINDNSADFIESLVDAKQAKELNEFRDGRLLTRGP 364
DB 304 QEEVSKLTVGDPEDNADITPVINDNSADFIWGLIEDAQEKQAALTPIKREGNLLMPVLE 363
YY 365 DHTYLMKLAEEBPFGPIIPIRVKDAEVAIAANKSDFGLOSSVFTDFOKAFDIANKL 424
DB 364 DHTYLMKLAEEBPFGPIIPIRVKDAEVAIAANKSDFGLOSSVFTDFOKAFDIANKL 423
YY 425 EVGTVINNKTKRGDPNPFPLGLKSGAGVQGIKYSIEAMTNVKSIVLDMK 475
DB 424 EVGTVINNKTKRGDPNPFPLGLKSGAGVQGIKYSIEAMTNVKSIVLDMK 474

RESULT 6

ADK47293 standard; protein; 474 AA.

XX ADK47293;
XX 20-MAY-2004 (first entry)
XX Streptococcus pneumoniae protein, Seq ID No 3808.
XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX Streptococcus pneumoniae.
XX US6699703-B1.
XX 02-MAR-2004.
XX 26-MAY-2000; 2000US-00583110.
XX 02-JUL-1997; 97US-0051553P.
XX 12-MAY-1998; 98US-0085131P.
XX 30-JUN-1998; 98US-00107433.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX WPI; 2004-212399/20.
XX N-PDB; ADK4632.

XX New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.

XX Disclosure; SEQ ID NO 3808; 301dp; English.

CC The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as S. pneumoniae infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.

XX Sequence 474 AA;

Query Match 75.6%; Score 1808.5; DB 8; Length 474;
Best Local Similarity 73.7%; Pred. No. 3.6e-142;
Matches 350; Conservative 59; Mismatches 65; Indels 1; Gaps 1;

YY 1 LTRKYONVNGEKKSVNOIELISPIDSSLGFPVPMTEEVDMKAGREAPAMALTYER 60
DB 1 LTRKYONVNGEKKSVNOIELISPIDSSLGFPVPMTEEVDMKAGREAPAMALTYER 59
YY 61 YVERAQLYKADIIERDKERIATVLAKEISKAVNASVTEVTRADLIRAAEGIRLST 120
DB 60 YVERAQLYKADIIERDKERIATVLAKEISKAVNASVTEVTRADLIRAAEGIRLST 119
YY 121 SDEEGKMDASTGHKLAVIRROPVGIATAPVNPVNLSSKIAAPALIGNVVMFKPPT 180
DB 120 QMEGGFPAASKNKLAVRRBPVGIATAPVNPVNLSSKIAAPALIGNVVMFKPPT 179
YY 181 QSSVSGVLAAPAGAGCPAGVFNITGRSGEIGDYIVHEBEVNFINTGSTPVGQRIG 240
DB 180 QSSVSGVLAAPAGAGCPAGVFNITGRSGEIGDYIVHEBEVNFINTGSTPVGQRIG 239
YY 241 LAGMPRIMLELGGKAGIATADADLNAKQIVAGAYDYGQRCTAIKRVLVVEVADL 300
DB 240 LAGMPRIMLELGGKAGIATADADLNAKQIVAGAYDYGQRCTAIKRVLVVEVADL 299
YY 301 AEKISENVAKLSVQDPEDNATVPVINDNSADFIESLVDAKQAKELNEFRDGRLLT 360
DB 300 ATTLQEEVSKLTVGDPEDNADITPVINDNSADFIWGLIEDAQEKQAALTPIKREGNLLM 359
YY 361 PGLPDRVTLDMKLAEEBPFGPIIPIRVKDAEVAIAANKSDFGLOSSVFTDFOKAFDI 420
DB 360 PGLPDRVTLDMKLAEEBPFGPIIPIRVKDAEVAIAANKSDFGLOSSVFTDFOKAFDI 419
YY 421 ANKLEVGTVHINKTKRGDPNPFPLGLKSGAGVQGIKYSIEAMTNVKSIVLDMK 475
DB 420 ANKLEVGTVHINKTKRGDPNPFPLGLKSGAGVQGIKYSIEAMTNVKSIVLDMK 474

RESULT 7

ABU46050 standard; protein; 474 AA.

XX ABU46050;
XX 19-JUN-2003 (first entry)
XX Protein encoded by prokaryotic essential gene #11577.
XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX Streptococcus pneumoniae.
XX MO20027183-A2.

PD 03-OCT-2002.
 XX 21-MAR-2002; 2002WC-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA49920.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 73974; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_poc_sequences
 XX
 SQ Sequence 474 AA;
 Query Match 75.3%; Score 1802; DB 6; Length 474;
 Best Local Similarity 73.7%; Pred. No. 1.3e-141;
 Matches 347; Conservative 58; Mismatches 66; Indels 0; Gaps 0;
 QY 5 YONVYVNGEKSSVNOIETLSPIDSSIGFVPANTREEVHAMKAGREALPAMALTVYR 64
 DB 4 YONLVNGKKSSSEQELTITSPINOEBLGTVPANTQTEADBEAMQAAALPAMALSAIR 63
 QY 65 AQTLEKADIIERDEKEIATVLAKETISKAYMSVTEVPTADLIRYAAEEGRLSTADE 124
 DB 64 AAYIHNTAAILREDEKEIGITLAKEVAKGIGKAAIGVVTADLIRYAAEEGRITGOAME 123
 QY 125 GGMMDASTGHTKAVIRROVGVTLATAPNYPNLSGSKTAPLLIGNTVMFPPQGVV 184
 DB 124 GGGFEALSKNKLAVARREPVGLTALAPNYPNLSASKTAPLLAGNTVMFKPPQGS 183

QY 185 SGLVLAKAFAEAGLPAGVFNITTGSGSEIGDYIHEHEVNFINTGSPVGORIGKLAGM 244
 DB 184 SGLLAKAFEEAGLPAGVFNITTGSGSEIGDYIHEHEVNFINTGSPVIGERIGRLAGM 243
 QY 245 RPIMLEGKQDAGIVLADADIDNNAKQIVAGAYDYSGORCTRAIKRVLVVEVADELAKEI 304
 DB 244 RPIMLEGKQDAAALVLEADIDLEHAKEQIVAGAFSYSGORCTRAIKRVLVLESVADKLATLL 303
 QY 305 SENVAKLSVSGDPEDNATVPYIDNNSADFISSLVVDAAOKAKXELNEKRGRLITPGLF 364
 DB 304 QEEVSKLTIVGDPEDNADITPVYIDNNSADFIWGLLEDAQEKRAQALTPIKRBSNLTMPVLV 363
 QY 365 DHVTIDMKLAMEEPGPILPIIRVDAEAAVANKSDFGQSSVFTBDPOKAPDIANKL 424
 DB 364 DQVTDKMVAWEPPGPVLPPIIRVASVEBALAFANESFGLQSSVFTNDPKAFPIANKL 423
 QY 425 EYGVTHINKTGRGPDNFPFLGLKSGAGVQGIKYSIEMTNVKSIVIDMK 475
 DB 424 EYGVTHINKTQGRGPDNFPFLGVKSGAGVQGIKYSIEMTNVKSIVPDKV 474
 RESULT 8
 AB044392
 ID AB044392 standard; protein; 475 AA.
 XX
 XX AB044392;
 AC 19-JUN-2003 (first entry)
 DT
 XX Protein encoded by Prokaryotic essential gene #29919.
 DE
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW
 XX Streptococcus mutans.
 OS
 XX WO200277183-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 21-MAR-2002; 2002WC-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA48262.
 DR
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 72316; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (9)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from Wipo at
CC ftp.wipo.int/pub/published_pct_sequences
CC
XX
SQ Sequence 475 AA;

Query Match 71.0%; Score 1700; DB 6; Length 475;
Best Local Similarity 66.9%; Pred. No. 4, 2e-133;
Matches 318; Conservative 75; Mismatches 82; Indels 0; Gaps 0;

QY 1 LTKYQYNYNGEWSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPAAWALT 60
DB 1 MTKQYQYNYNGEWSVNOIEILSPIDSSLGFPVPMTEBEVDHAKAGREALPAAWALT 60
QY 61 YVERAQLHAKADIRDXEATVLAKEISKAYNASVTEVATDILRYAEBEGRIST 120
DB 61 YVERAQLHAKADIRDXEATVLAKEISKAYNASVTEVATDILRYAEBEGRIST 120
QY 121 SAEDEGKMDASTGHKLAIVRROPVGIIVLAIPANYVNISSGKIAALIGANVWFKPPT 180
DB 121 EYDEGGSFPAASKKIAVVRREBVGIVLAISPNTVNLGSKIAALIGANVWFKPPT 180
QY 181 QGSVSGVLAKAPAEAGLPAGVNTITGSGEIGDYIVHEEVNFTNGTSPVGGRIK 240
DB 181 QGSISGLLAAEAPAGLPAGVNTITGSGEIGDYIVHEEVNFTNGTSPVGGRIK 240
QY 241 LAGMRPIMELGSKDAGIYLAADLDMNAKOIYAGVYSGORCTIAKRYLVVEEYADL 300
DB 241 LAGMRPIMELGSKDAGIYLAADLDMNAKOIYAGVYSGORCTIAKRYLVVEEYADL 300
QY 301 AEKISNVAKLSVDEPDNATVPVTDNSADPISLVVDARQKAKEINEFRDGRILT 360
DB 301 VEKIREKVALITIGNEBDDADITPLIDTYSADIVBGLINDANDKGAALFEIKREGNLIC 360
QY 361 PGHFDHTLDMKLAWEPEFGPIPIIRVDAEBAVIAKSPDGLSSVETDFOKAPDI 420
DB 361 PILEFDVTTDMRLAWEPEFGPIPIIRVSVBEAITSKSEVGLQASIFTDFFPAFBI 420
QY 421 ANKLEVGTVHINNKTGRGPDNPFPLGKSGAGVQIRYSIEAMTVKSIIVLDMK 475
DB 421 AEOLBVGTVHINNKTGRGPDNPFPLGKSGAGIQCIVKYSIEAMTVKSVVEEDIK 475

RESULT 9
AAU37723
ID AAU37723 standard; protein; 439 AA.

AC AAU37723;
XX
XX
DT 14-FEB-2002 (first entry)
DB Streptococcus pneumoniae cellular proliferation protein #152.
XX
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.

OS Streptococcus pneumoniae.

XX PN WO200170955-A2.

XX 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2001; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GD;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AAS55582.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 13316; 511p; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from Wipo at
CC ftp.wipo.int/pub/published_pct_sequences
CC
XX
SQ Sequence 439 AA;

Query Match 71.0%; Score 1699; DB 4; Length 439;
Best Local Similarity 74.5%; Pred. No. 4, 5e-133;
Matches 327; Conservative 55; Mismatches 57; Indels 0; Gaps 0;

QY 37 MTRBEVDHAKAGREALPAAWALTVERAQLHAKADIRDXEATVLAKEISKAYNA 96
DB 1 MTRBEVDHAKAGREALPAAWALTVERAQLHAKADIRDXEATVLAKEISKAYNA 96
QY 97 SVEVVRATADILRYAEBEGRISTSADEGKMDASTGHKLAIVRROPVGIIVLAIPANY 156
DB 97 SVEVVRATADILRYAEBEGRISTSADEGKMDASTGHKLAIVRROPVGIIVLAIPANY 156
QY 61 AIGEVVRINDILRYAEBEGRITGQMEGGEAASKKIAVVRREBVGIVLAIPANY 120
DB 61 AIGEVVRINDILRYAEBEGRITGQMEGGEAASKKIAVVRREBVGIVLAIPANY 120
QY 121 VNLSSKIAIPALIGNVVWFKPPTGSGVGLVLAFAEAGLPAGVNTITGSGEIGDY 180
DB 121 VNLSSKIAIPALIGNVVWFKPPTGSGVGLVLAFAEAGLPAGVNTITGSGEIGDY 180
QY 217 IVEHEVNFTNGTSPVGGRIKLAGMRPIMELGSKDAGIYLAADLDMNAKOIYVGA 276
DB 217 IVEHEVNFTNGTSPVGGRIKLAGMRPIMELGSKDAGIYLAADLDMNAKOIYVGA 276
QY 277 YDYGORCTIAKRYLVVEEYADLAEKISNVAKLSVDEPDNATVPVTDNSADPIS 336
DB 277 YDYGORCTIAKRYLVVEEYADLAEKISNVAKLSVDEPDNATVPVTDNSADPIS 336
QY 241 PYSQGRCTIAKRYLVVEEYADLAEKISNVAKLSVDEPDNATVPVTDNSADPIS 300
DB 241 PYSQGRCTIAKRYLVVEEYADLAEKISNVAKLSVDEPDNATVPVTDNSADPIS 300

Qy	337	L V I D A R O G A K E I N F E K D G R I L T G L P H Y I L D M K L A E E P R G I L P I I R Y K D A E A Y A	396
Db	301	L I E D Q E K E A Q M L T P I K K E G N I L M P V L F Q V T K D M K V A M E E P G P V L P I I R V A S V E A T A	360
Qy	397	I A N K S D F G I Q S S V F P T R D P F O K A F E D I A N K L E V G I V H I N N K T G R G P D N P F F I G L K G S G A G V O G	456
Db	361	F A N E S E F G I Q S S V F I N D F P K A F E I A K E L E V G I V H I N N K T Q R G P D N P F F I G Y K S G A G V O G	420
Qy	457	I R Y S I E A M T N V K S I Y L D M K	475
Db	421	I K Y S I E A M T N V K S I V E D V K	439
RESULT 10			
ADRS95169	ID	ADRS95169 standard; protein; 436 AA.	
XX	AC	ADRS95169;	
XX	DT	16-DEC-2004 (first entry)	
XX	DE	Novel S. pneumoniae protein sequence, SEQ ID 3804.	
XX	KW	Meningitis; bacteraemia; pneumonia; otitis media; vaccine;	
XX	KW	bacterial infection.	
XX	OS	Streptococcus pneumoniae.	
XX	PN	US6800744-B1.	
XX	PD	05-OCT-2004.	
XX	PE	30-JUN-1998; 98US-00107433.	
XX	PR	02-JUN-1997; 97US-0051553P.	
XX	PR	12-MAY-1998; 98US-0085131P.	
XX	PA	(GENO-) GENOME THERAPEUTICS CORP.	
XX	PI	Doucette-Stamm LA, Bush D;	
XX	DR	WPI; 2004-697205/68.	
XX	DR	N-PSDB; ADR92566.	
XX	PT	New isolated nucleic acid encoding a Streptococcus pneumoniae	
XX	PT	polypeptide, useful for diagnosing, preventing and/or treating	
XX	PT	pathological conditions resulting from the bacterial infection.	
XX	PS	Disclosure; SEQ ID NO 3804; 151pp; English.	
XX	CC	The invention relates to an isolated nucleic acid comprising a sequence	
XX	CC	encoding a Streptococcus pneumoniae ADR91366polypeptide, or its	
XX	CC	fragments, with any of 9 fully defined sequences (appearing as ADR94308,	
XX	CC	ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,	
XX	CC	ADR96079) or any of the fully defined sequences appearing as ADR91705,	
XX	CC	ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR9366, ADR9650 or	
XX	CC	ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide	
XX	CC	sequences, or at least 40, 60 or 300 consecutive nucleotides, which is	
XX	CC	hybridisable under high stringency conditions to the nucleotide sequence.	
XX	CC	The nucleic acids and proteins are chosen from 5206 disclosed sequences.	
XX	CC	Also included are a recombinant expression vector comprising the isolated	
XX	CC	nucleic acid cited above operably linked to a transcription regulatory	
XX	CC	element, a cell comprising the recombinant expression vector and a probe	
XX	CC	comprising at least 20 consecutive nucleotides of the nucleotide	
XX	CC	sequences as cited above. The methods and compositions of the present	
XX	CC	invention are useful for the diagnosis, prevention and/or treatment of	
XX	CC	pathological conditions resulting from bacterial infection by	
XX	CC	Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and	
XX	CC	otitis media. The present sequence is one of the 2603 disclosed S.	
XX	CC	pneumoniae protein sequences. Note: The sequence data for this patent did	
XX	CC	not form part of the printed specification, but was obtained in	
XX	CC	electronic format directly from USPTO at	

CC	segdata.uspfo.gov/sequence.html?docID=6600474AB1.
XX	Sequence 436 AA;
XX	Query Match 70.7%; Score 1692; DB 8; Length 436;
XX	Best Local Similarity 74.9%; Pred. No. 1.7e-132;
XX	Matches 326; Conservative 53; Mismatches 56; Indels 0; Gaps 0
QY	41 EVDHAKAGREALPAWALITVERAQYIHKADIIERDKIEIATVLAKESIKAYNASYTE 1000
DB	2 EADGQQAARALPAARALSAIERAAYLHKTAALIERDKKIGTILAKETAKGKAIGE 61
QY	101 VVPTADLIRYAABEGRLSTADSEGRKMDASTGHKLAVIRROQPIVLAIAAPVNPVLS 166
DB	62 VVPTADLIRYAABEGRLITGQMEGGFEASKNKLAVVREPVGIVLAIAPEVNPVLS 121
QY	161 GSKIAIPALIGSNVMEKPPQTQSGVLIAKAFAPAGVFTITGRGSEIGDYIYEH 220
DB	122 ASKIAIPALIGSNVMEKPPQTQSGIGLLAKAFEBAGIPAVVFTITGRGSEIGDYIIEH 181
QY	221 EEVNFINFTGSPVVGORIGKLAKMFPIMLEIGSGDAGIVLADADLDNNAKOIVAGAYDS 280
DB	182 KEVNFINFTGSPFIGERIGRLAKMFPIMLEIGSKDPAALVLEDDADLEHAAKOIVAGAPSYS 241
QY	281 GORCTIKIKVUVEVEADLEAKIESBNAXISVGDPPNATVTFVIDINDSNADFIESIYVD 340
DB	242 GORCTIKIKVUVEVEADLATLLDEEVSCKLVGDPPFNADITVIDINDSNADFTWGLIED 301
QY	341 AROKGAKEINFEPRDRLITPGLPDHVTLDMLKAMEBPFGDILPIIRVKDAEVAIAANK 400
DB	302 AORKEKQALTPIKREBNLMPVLPFOVTKDKKVAKEEPFGVLDIIRVASVEEALAPANE 361
QY	401 SDFGLQSVFTRDFOKAPDIANKLEVGTVHINNTKGRGPNFPFLGKSGAGVQGIKYS 460
DB	362 SERGLQSVFTRDFOKAFETIAKLEVGTVHINNTKGRGPNFPFLGKSGAGVQGIKYS 421
QY	461 IEAMTVKSIIVLDMK 475
DB	422 IEAMTVKSIIVPDVK 436
XX	RESULT 11
XX	ABP27644
XX	ID ABP27644 standard; protein; 475 AA.
XX	AC
XX	ABP27644;
XX	DT 02-JUL-2002 (first entry)
XX	DE Streptococcus polypeptide SBQ ID NO 4464.
KW	Streptococcus GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM	group A streptococcus; Streptococcus pyogenes; antibacterial;
KM	antiinflammatory; infection; vaccine; meningitis; gene therapy.
OS	Streptococcus pyogenes.
XX	XX
PM	W0200234771-A2.
XX	XX
PD	02-MAY-2002.
XX	XX
PF	29-OCT-2001; 2001WO-GB004789.
XX	XX
PR	27-OCT-2000; 2000GB-00026333.
PR	24-NOV-2000; 2000GB-00028727.
PR	07-MAR-2001; 2001GB-00005640.
XX	XX
PA	(CHIR-) CHIRON SPA.
PA	(GENO-) INST GENOMIC RES.
XX	XX
PI	Telford J, Maignani V, Margarit Y Rosl, Grandi G, Frazer C;
XX	Tetteijn H;

DR MPI; 2002-352536/38.
DR N-PSDB; ABN68275.
PT New Streptococcus protein for the treatment or prevention of infection or
XX disease caused by Streptococcus bacteria, such as meningitis, and for
XX detecting a compound that binds to the protein.
PS Claim 1; Page 3607; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and anti-inflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
XX Sequence 475 AA;
SQ
Query Match 68.6%; Score 1641; DB 5; Length 475;
Best Local Similarity 65.7%; Pred. No. 3.5e-128;
Matches 312; Conservative 76; Mismatches 87; Indels 0; Gaps 0;
1 LTKXYONYNGEMKSSVNOIEILSPIDBSLGFVPMTRBEVDHAMKAGREALPMAALT 60
1 LAKQYKNLVNGEMKSSVNOIEILSPIDBSLGFVPMTRBEVDHAMKAGREALPMAALT 60
61 YVERAOYLHKADIIERDKSEIATVLAKEISAKYNAVSTEVRADLIRYAEGRIST 120
61 YVERAOYLHKADIIERDKSEIATVLAKEISAKYNAVSTEVRADLIRYAEGRIST 120
61 YVERAOYLHKADIIERDKSEIATVLAKEISAKYNAVSTEVRADLIRYAEGRIST 120
121 SAEDEGKMDASTGHKLAIVRRPVGIVLAIPYVNVLSGSKIAPALIGVNVMPKPT 180
121 EYLBEGSEFAASKKKAIVRRPVGIVLAIPYVNVLSGSKIAPALIGVNVMPKPT 180
181 QGSVGLVLAFAEAGLPAGVFNITGRGSEIGDIYVEHEEVNFINTGSTPVGQRIK 240
181 QGSISGLLAAFAEAGLPAGVFNITGRGSEIGDIYVEHEEVNFINTGSTPVGQRIK 240
241 LAGMRPIMELGKGKAGIYLAADADLADNAKQIVAGAYDYSQGRCTAIRKVLVEYADEL 300
241 LAGMRPIMELGKGKAGIYLAADADLADNAKQIVAGAYDYSQGRCTAIRKVLVEYADEL 300
301 AEIKTIVKLSVGPEDDADITPILDTISADPVEGLIKDADTKGATALTAFNRBGNLIS 360
301 AEIKTIVKLSVGPEDDADITPILDTISADPVEGLIKDADTKGATALTAFNRBGNLIS 360
361 PCLEFHTVLTMDKLAWEPEGPILPIIRYKDAEEAVAIANKSDGLQSSVFTDFOKAFPI 420
361 PCLEFHTVLTMDKLAWEPEGPILPIIRYKDAEEAVAIANKSDGLQSSVFTDFOKAFPI 420
421 ANKLEVGTVHINKTGRGPDNPFPLGLKSGSAGVQGIKIRYSIEAMTVKSIIVLDMK 475
421 ANKLEVGTVHINKTGRGPDNPFPLGLKSGSAGVQGIKIRYSIEAMTVKSIIVLDMK 475
421 ANKLEVGTVHINKTGRGPDNPFPLGLKSGSAGVQGIKIRYSIEAMTVKSIIVLDMK 475
RESULT 12
ID ADR83939 standard; protein; 475 AA.
XX ADR83939;
AC ADR83939;
XX 02-DEC-2004 (first entry)
DT

XX
XX S. pyogenes hyperimmune system reactive antigen SpY1371.
DE hyperimmune serum reactive antigen; vaccine; anticaline.
XX hyperimmune serum reactive antigen; vaccine; anticaline.
XX Streptococcus pyogenes.
OS Streptococcus pyogenes.
XX
XX WO2004078907-A2.
XX
XX 16-SEP-2004.
XX
XX 02-MAR-2004; 2004WO-EP002087.
XX
XX 04-MAR-2003; 2003EP-00450061.
XX
XX (INTE-) INTERCELL AG.
XX
XX Meinke A, Nagy E, Winkler B, Gelbmann D;
PI MPI; 2004-653698/63.
DR N-PSDB; ADR83789.
XX
XX New isolated nucleic acid molecules encoding hyperimmune serum-reactive
PT antigens from Streptococcus pyogenes, useful for diagnosing, preventing
PT and treating S. pyogenes infections.
XX
XX Claim 11; SEQ ID NO 207; 145pp; English.
XX
XX This invention describes a novel nucleic acid molecule encoding a
CC hyperimmune serum reactive antigen or its fragment from Streptococcus
CC pyogenes. The nucleic acid molecule or hyperimmune serum-reactive antigen
CC or its fragment are useful for the manufacture of a pharmaceutical
CC preparation, especially a vaccine, against S. pyogenes infection. In
CC addition, the hyperimmune serum reactive antigen or fragment is used for
CC the isolation and/or purification and/or identification of an interaction
CC partner of the hyperimmune serum reactive antigen or its fragment, for
CC the generation of a peptide (e.g. anticalins) binding to the antigen or
CC fragment, or for the manufacture of a functional nucleic acid selected
CC from aptamers and spiegelmers. The nucleic acid molecule may also be used
CC for the manufacture of functional ribonucleic acids, such as ribozymes,
CC antisense nucleic acids and siRNA. ADR83733-ADR84189 represent S.
CC pyogenes hyperimmune serum reactive antigens, fragments and the encoding
CC polynucleotide described in the invention.
XX
XX Sequence 475 AA;
SQ
Query Match 68.6%; Score 1641; DB 8; Length 475;
Best Local Similarity 65.7%; Pred. No. 3.5e-128;
Matches 312; Conservative 76; Mismatches 87; Indels 0; Gaps 0;
1 LTKXYONYNGEMKSSVNOIEILSPIDBSLGFVPMTRBEVDHAMKAGREALPMAALT 60
1 LAKQYKNLVNGEMKSSVNOIEILSPIDBSLGFVPMTRBEVDHAMKAGREALPMAALT 60
61 YVERAOYLHKADIIERDKSEIATVLAKEISAKYNAVSTEVRADLIRYAEGRIST 120
61 YVERAOYLHKADIIERDKSEIATVLAKEISAKYNAVSTEVRADLIRYAEGRIST 120
61 YVERAOYLHKADIIERDKSEIATVLAKEISAKYNAVSTEVRADLIRYAEGRIST 120
121 SAEDEGKMDASTGHKLAIVRRPVGIVLAIPYVNVLSGSKIAPALIGVNVMPKPT 180
121 EYLBEGSEFAASKKKAIVRRPVGIVLAIPYVNVLSGSKIAPALIGVNVMPKPT 180
181 QGSVGLVLAFAEAGLPAGVFNITGRGSEIGDIYVEHEEVNFINTGSTPVGQRIK 240
181 QGSISGLLAAFAEAGLPAGVFNITGRGSEIGDIYVEHEEVNFINTGSTPVGQRIK 240
241 LAGMRPIMELGKGKAGIYLAADADLADNAKQIVAGAYDYSQGRCTAIRKVLVEYADEL 300
241 LAGMRPIMELGKGKAGIYLAADADLADNAKQIVAGAYDYSQGRCTAIRKVLVEYADEL 300
301 AEIKTIVKLSVGPEDDADITPILDTISADPVEGLIKDADTKGATALTAFNRBGNLIS 360
301 AEIKTIVKLSVGPEDDADITPILDTISADPVEGLIKDADTKGATALTAFNRBGNLIS 360

XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX MPI: 2003-029926/02.
 DR N-PSDB; ACA29489.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 53543; 1766bp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 486 AA:
 SO
 Query Match 58.4%; Score 1397; DB 6; Length 486;
 Best Local Similarity 56.9%; Pred. No. 8.8e-108;
 Matches 272; Conservative 80; Mismatches 122; Indels 4; Gaps 2;

QY 298 DELAEKISENVAKLGVDEPDNATVPTVDNSADPESLAVVDAROKAKELNEFRDGR 357
 DB 307 DKLVGKIKPLVEKLTIGNPMDEVITPLDNRKATDFVQGLVDDALHKAKLITGNVRKN 366
 QY 358 LITPGIFDHVTLDMKLAWEPPGPIPIIRVDAEBAVAIVANKSDGLQSVFTDPOKA 417
 DB 367 LFVPTLDDVNVNDMKIAMEPPGVPVPIIRVNDINQAIETANQSEVGLQSVFTSDIDKA 426
 QY 418 FDIANKLEVTYTHINNKTRGPDNPFPLGLKSGAGVQGIKRSIEMTVKSLVLDMK 475
 DB 427 FYIADKLGVTVQINNKTERGPDHPEPLGVKASGMGTQGVKYSIEMTRPKAVVAVNR 484
 RESULT 15
 ID ABU24815
 ABU24815 standard; protein; 486 AA.
 AC ABU24815;
 DT 19-UTN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #10342.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 OS Clostridium botulinum.
 PN W0200277183-A2.
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002MO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362899P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX MPI: 2003-029926/02.
 DR N-PSDB; ACA28685.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 52739; 1766bp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of


```

DB 251 GMLPWLLEIGKDAIVLSDANLETTAKSIVSAYSGRCAYRVLYMKVADLELVE 310
QY 303 KISENVAKISVGDPPDNATVTPVIDNSADPISLVDARQKAKELNEFKRDGRLITPG 362
DB 311 LVTKKKELEKLVGNPFDDVITPLIDRKADYVQTLIDDAIEKATLIVNKKRENIAMPT 370
QY 363 LFDHVTLMKLAKEEPFGPIPIIRKDAEENVALINKSGDFGQSSVFTPRDPOKAPDIAN 422
DB 371 LFDNVTADMKRIAEPPFGVPLPIIRKXSDAELELNREBYGLQSAVFTEHMDARYIAN 430
QY 423 KLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQIGIRYSIEAMTNVKSIVLDM 474
DB 431 KLDVGTVOYNNKREGRGPDHPFLGTSSGMDGQIGIRYSIEAMTRHKSIVLNL 482

RESULT 17
ADS27503
ID ADS27503 standard; protein; 472 AA.
XX
AC ADS27503;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #16536.
XX
KM Recombinant DNA construct; transformed plant; improved plant property;
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KM pathogen tolerance; pest tolerance; plant disease resistance;
KM cell cycle pathway modification; plant growth regulator;
KM homologous recombination; seed oil yield; protein yield; carbohydrate;
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KM bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 16536; 122pp; English.
XX
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transgenic plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

```

```

CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 472 AA;
XX
Query Match 55.8%; Score 1335; DB 8; Length 472;
Best Local Similarity 55.7%; Pred. No. 1.3e-102;
Matches 263; Conservative 80; Mismatches 127; Indels 2; Gaps 1;

QY 5 YGNVNGEMK--SSVNOJEIISPIDDSSLGFPVPMTRREVDHAKAGREALPAMALITVY 62
DB 1 YKFYINGEWRSSSEETIEIPSPYHBEVYGVOALTRGEVDALSAKEAQSMABASIQ 60
QY 63 ERAQYIHKAAADIERDKESIATVLAKEISKAYNASVTEVYRTADLIRYAEGRILSTSA 122
DB 61 DRAKYLYKRADELVMQDEIADIIMKEVGKGYKDAKKEVYRTADLIRYIEALMHGGS 120
QY 123 DEGGKMDASTGKLAIVIRROPVGIYLAIPVNPVNLSSGKTAPALIGNVMEKPPVQ 182
DB 121 NMGDSFPGGTSGKLAIIORAPGVVLAIPFVYPVNLSPAUKLAPALIMGNNAVIFRPATQ 180
QY 183 SVSGVLAKAFPAEALPGVFNTIIRGSEIGDYVEHEVNFIPQTSPIVQGIIGKLA 242
DB 181 AISGIKWEALHAKGLPGSLVNAVATGKSGVIGDYVEHEGIMMVSTPGTNTGKHLAKKA 240
QY 243 GMRPIMLEIGKDAIVLADADLDNAKQIVAGAYDSGORCTAIKRVLVVEEVADELAE 302
DB 241 SMIPVLLEIGKDPRIAREDDLOPAANHIVSGAYSQGRCTAIKRVLVHENVADDELVD 300
QY 303 KISENVAKISVGDPPDNATVTPVIDNSADPISLVDARQKAKELNEFKRDGRLITPG 362
DB 301 LKQAVAEISVSGPQDSTVPLIDKSDAFQGLVDVAVEKATVIGNKKERNLIYPT 360
QY 363 LFDHVTLMKLAKEEPFGPIPIIRKDAEENVALINKSGDFGQSSVFTPRDPOKAPDIAN 422
DB 361 LIDHVTLEKVAKEEPFGPIPIIRVSSDEQAIETIANSEFGLQSAVFPKDINKAPALAN 420
QY 423 KLEVGTVHINNKTGRGPDNPFPLGLKSGAGVQIGIRYSIEAMTNVKSIVLDM 474
DB 421 KIETSGVOINGRTERGPDHPFPIGVKSGMGAGIRKLSLSTREKIVLNL 472

RESULT 18
ADS27631
ID ADS27631 standard; protein; 473 AA.
XX
AC ADS27631;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #16664.
XX
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX

```


AC ABM74113;
XX
DT 17-OCT-2003. (first entry)
XX
DE DNA clone originating in barley containing SNP sequence #523.
XX
KM Barley; single nucleotide polymorphism, SNP; genotype-phenotype analysis.
OS Hordeum vulgare.
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-1B005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
XX
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284bp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences
XX
SQ Sequence 496 AA;

Query Match 48.9%; Score 1169.5; DB 7; Length 496;
Best Local Similarity 49.8%; Pred. No. 9.1e-89;
Matches 238; Conservative 84; Mismatches 143; Indels 13; Gaps 5;

QY 5 YONYVNGEKKSSVN--QIRILSPIDDSIGFVPMTRREVDHMKRGREALPMALTYV 62
DB 16 YKTYADGEMRPASGKTVAIVNPTROTQYRVQACTQEEENKWKMDAKVAKQKMATPTLM 75
QY 63 ERAQYLHKAAADIERDKKEIATFLAKEISKAVNASTEVETADLIRYAAEGIRLSTSA 122
DB 76 KKAELHLKAAALIKHKETPIAECLVKEIKKPADAVSEVVRSGDLVSYTAEEGVRLIG- 134
QY 123 DEEGKM--DASTG--HKLAVIRQPGVIALAIYNTYPVNLSGSKIAPALIGNVNMF 176
DB 135 ---GKLVSDSPFGERNRNYKCLSSKVPGLGVLAIPFPNYPVNLAIVSKIGPALIAGNSLVY 191
QY 177 KPRTGSGVGLVYAKAPAEAGLPAVENVITTGSGSEIGDYIVHEEYVNFNFTGSPVQO 236
DB 192 KPRTQAVVAALHNVHCFHLNAGFPKGLISCTGKSGSEIGDFTLHHPGVNCLSTFGG-DTGI 250
QY 237 RIGKLAKMPIMLELGKDGAGIVLADADLDNNAKOIVAGAYDYSGORCTAIRKVLVEEV 296
DB 251 AISKAGWVPLQWELGSKDCACTVLEDDADLDVLAANIVKGFSTSGRCRAVAKVLLMEV 310

QY 297 ADELAEKISENVAKLVSGDPFDNATVTPYIDNSADFIESTLVVDAROKGAKELNEKPRDG 356
DB 311 ADAVEKVNAKAKLKVGPPEDDCITPVYTSANFIEGLVMDAKEKATTCQETRRRG 370
QY 357 RLITRGLFDHVTLLDMKLAMEBPFGPILPIIRVKDAEAVAIANKSDPGLQSSVFTTRDPK 416
DB 371 NLIWPLLDHVPDRIANWEEPFGLPVLRINSVVEEGIHHCNANFGLQGVCFTRDINK 430
QY 417 AFDIANKLEVGTVHINNKTRGPDNPPIGLKSGAGVCGIYRSIEAMTVKSYILDM 474
DB 431 AIMISDAMETGVQINSAPARGPDHFPFQGLKDSIGISQGITNSINMTKVKSVTINTL 488

RESULT 22
AAG32607
ID AAG32607 standard; protein; 496 AA.
XX
AC AAG32607;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 39371.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-01231825P.
PR 05-MAR-1999; 99US-01231808P.
PR 09-MAR-1999; 99US-01235488P.
PR 23-MAR-1999; 99US-01257888P.
PR 25-MAR-1999; 99US-01262649P.
PR 29-MAR-1999; 99US-01267859P.
PR 01-APR-1999; 99US-01274628P.
PR 06-APR-1999; 99US-01282349P.
PR 08-APR-1999; 99US-01287149P.
PR 16-APR-1999; 99US-01298459P.
PR 19-APR-1999; 99US-01300779P.
PR 21-APR-1999; 99US-01304499P.
PR 23-APR-1999; 99US-01305109P.
PR 23-APR-1999; 99US-01308919P.
PR 28-APR-1999; 99US-01314499P.
PR 30-APR-1999; 99US-01320489P.
PR 30-APR-1999; 99US-01324079P.
PR 04-MAY-1999; 99US-01324849P.
PR 05-MAY-1999; 99US-01324859P.
PR 06-MAY-1999; 99US-01324869P.
PR 06-MAY-1999; 99US-01324879P.
PR 07-MAY-1999; 99US-01328639P.
PR 11-MAY-1999; 99US-01332569P.
PR 14-MAY-1999; 99US-01334218P.
PR 14-MAY-1999; 99US-01342199P.
PR 14-MAY-1999; 99US-01342219P.
PR 14-MAY-1999; 99US-01343709P.
PR 18-MAY-1999; 99US-01347689P.
PR 19-MAY-1999; 99US-01349419P.
PR 20-MAY-1999; 99US-01351249P.
PR 21-MAY-1999; 99US-01353539P.
PR 24-MAY-1999; 99US-01356299P.
PR 25-MAY-1999; 99US-01360219P.
PR 27-MAY-1999; 99US-01363929P.
PR 28-MAY-1999; 99US-01367829P.
PR 01-JUN-1999; 99US-01372229P.
PR 03-JUN-1999; 99US-01375289P.
PR 04-JUN-1999; 99US-01375029P.

PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148311P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0153759P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155148P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0158665P.
PR 07-OCT-1999; 99US-0158682P.
PR 08-OCT-1999; 99US-0158623P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161921P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 48.8%; Score 1168.5; DB 3; Length 496;
Best Local Similarity 50.3%; Pred. No. 1.1e-88;
Matches 240; Conservative 84; Mismatches 142; Indels 11; Gaps 5;

29-APR-2004.
20-OCT-2003, 2003WO-EP011658.
18-OCT-2002, 2002EP-00079408.
(CROP-) CROPPDESIGN NV.
Inze D, De Veylder L, Vlieghe K;
WPI: 2004-348466/32.
N-PSDB; ADN73354.
Altering plant characteristics, useful for producing plants for enzyme or pharmaceutical production comprises modifying in a plant, expression of one or more nucleic acids and/or modifying level or activity of one or more proteins.
Claim 1, SEQ ID NO 1250, 134pp, English.

This invention relates to a novel method for altering one or more plant characteristics. Specifically, it refers to identifying genes that are up- or down-regulated in transgenic plants overexpressing the heterodimeric E2Fs/Dpa transcription factor of *Arabidopsis* and using these sequences to alter plant characteristics accordingly. The present invention describes generating transgenic plants for the production of growth regulators, enzymes, therapeutics, pharmaceuticals and animal feed products, where the altered plant characteristics are selected from increased yield or biomass, enhanced survival capacity, stress tolerance, plant architecture or physiology, altered lipid mobilization, biochemistry, signal transduction, storage lipid mobilization and/or altered photosynthesis, each relative to the corresponding wild type plants. Accordingly, these sequences can also be useful as positive or negative selectable markers during transformation of cells or tissues. The identified genes play a role in a variety of biological processes such as DNA replication, cell wall biosynthesis, nitrogen and/or carbon metabolism or they function as transcription factors. This polypeptide sequence is that cress protein expressed by a gene upregulated 1.3 fold or more in plants overexpressing the E2pa/Dpa transcription factor, given in an exemplification of the invention.

Sequence 496 AA;

Query Match	48.8%;	Score 1168.5;	DB 8;	Length 496;
Best Local Similarity	50.3%;	Pred. No. 1.1e-88;		
Matches 240;	Conservative 84;	Mismatches 142;	Indels 11;	Gaps 5

```

Oy 5 0N9NYNGEKK--SSVNOJ01111PIDDSSIGFUPAMIRREEDYHAMKRGREULPMAWLTJY 62
Db 16 YKYVDGEEKTSSGGSKVAINMPARTRKTYQYOACT0EBEYNAVMELAKSHQSMATPIW 75
Oy 63 BRAOYLHKAADIIERDKBEIATVLAKESISKAVNASVTEVTRADLIIRYAAEBGIRLSTSA 122
Db 76 KRAELHTRKAAAILTKONKAPMAESLVKEIAKPAKADSTVEVRSODLISYCAEBGVRL--L 132
Oy 123 DEGGCM--DASGCH--KLAVIRROPVGLVLAIPRNYPVNLSGSLAIPALLIGSNVMMK 177
Db 133 GEGKRLSDSPFGNDRYTKCLTSLKPLGVALLIPPNYVPVNLAVSKIAIAPLLANGSELVWK 192
Oy 178 PPTQGSVSGVLVAKAFABEAGLPAGVFNITTRGSGEIGDYIVHEHEVNFNFNGSTPEVQOR 237
Db 193 PPTQGAVSGCLHNVHCFHLGAPFKGLSCLTGGSGELGDP LTMHPRVNCISFPGC-DTGIS 251
Oy 238 IGLLAGRPIMLELGSKDAGIYVLADADLDNNAKQIYAGAVDYSGORCTAIKRYLVVEYA 297
Db 252 ISKKQGMGMLPLOWELGSKDACYLADADLDLVASNIKKGFSSYGORCTAVKVLVMEYSVA 311
Oy 298 DELAEKISENVAKLSTGVDPDNATVTPVLDONNSADPITRESIYVDAR0KGAKEINFPRDDR 357
Db 312 DELVEKYAKVAKVATLYGPPBENSDDITAVVSESSANTIEGLVMDAKKGAFTFOCEYKREBN 371
Oy 358 LITPGLFDHVTLMDKLAMEEFPQILPIIRVADAEBAVALANKSDFGLOSSVFTRDFOKA 417

```

Db	372	LIPWILLDNVRPDKMAIEEPFGPVVPLVLRINSVEBGINHCNANSNFGLOGCVTKRINKA	431
Qy	418	FDIANKEVEGVTHINNKTKRGCPDNEPFYGLKSGGAGVOGIRYSIEAMTVKSLVLDM	474
Db	432	ILISDAMEGTQVINSAPARGDHPFPQGLKDSGIGSGCVTNSINMTKVKTTVINL	488
RESULT	25		
ID	AAG05753		
XX	AAG05753	standard; protein; 496 AA.	
AC	AAG05753;		
DT	17-OCT-2000	(first entry)	
XX			
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 2276.	
XX			
KM	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
XX	termination sequence.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000;	2000EP-00301439.	
XX			
PR	25-FEB-1999;	99US-0121825P.	
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	16-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	23-APR-1999;	99US-0130891P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	07-MAY-1999;	99US-0132487P.	
PR	11-MAY-1999;	99US-0132863P.	
PR	14-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	14-MAY-1999;	99US-0134370P.	
PR	18-MAY-1999;	99US-0134768P.	
PR	19-MAY-1999;	99US-0134941P.	
PR	20-MAY-1999;	99US-0135124P.	
PR	21-MAY-1999;	99US-0135353P.	
PR	24-MAY-1999;	99US-0135629P.	
PR	25-MAY-1999;	99US-0136032P.	
PR	27-MAY-1999;	99US-0136392P.	
PR	28-MAY-1999;	99US-0136782P.	
PR	01-JUN-1999;	99US-0137222P.	
PR	03-JUN-1999;	99US-0137528P.	
PR	04-JUN-1999;	99US-0137502P.	
PR	07-JUN-1999;	99US-0137724P.	
PR	08-JUN-1999;	99US-0138094P.	
PR	10-JUN-1999;	99US-0138540P.	
PR	10-JUN-1999;	99US-0138847P.	
PR	14-JUN-1999;	99US-0139119P.	
PR	16-JUN-1999;	99US-0139452P.	

PR 09-MAR-1999: 99US-0123548P.
PR 23-MAR-1999: 99US-0125788P.
PR 25-MAR-1999: 99US-0126264P.
PR 29-MAR-1999: 99US-0126785P.
PR 01-APR-1999: 99US-0127462P.
PR 06-APR-1999: 99US-0128234P.
PR 08-APR-1999: 99US-0128714P.
PR 16-APR-1999: 99US-0129845P.
PR 19-APR-1999: 99US-0130077P.
PR 21-APR-1999: 99US-0130449P.
PR 23-APR-1999: 99US-0130510P.
PR 28-APR-1999: 99US-0131449P.
PR 30-APR-1999: 99US-0132048P.
PR 30-APR-1999: 99US-0132407P.
PR 04-MAY-1999: 99US-0132484P.
PR 05-MAY-1999: 99US-0132485P.
PR 06-MAY-1999: 99US-0132486P.
PR 06-MAY-1999: 99US-0132487P.
PR 07-MAY-1999: 99US-0132863P.
PR 11-MAY-1999: 99US-0134256P.
PR 14-MAY-1999: 99US-0134218P.
PR 14-MAY-1999: 99US-0134219P.
PR 14-MAY-1999: 99US-0134221P.
PR 14-MAY-1999: 99US-0134370P.
PR 18-MAY-1999: 99US-0134768P.
PR 19-MAY-1999: 99US-0134941P.
PR 20-MAY-1999: 99US-0135124P.
PR 21-MAY-1999: 99US-0135353P.
PR 24-MAY-1999: 99US-0135629P.
PR 25-MAY-1999: 99US-0136021P.
PR 27-MAY-1999: 99US-0136392P.
PR 28-MAY-1999: 99US-0136782P.
PR 01-JUN-1999: 99US-0137522P.
PR 03-JUN-1999: 99US-0137528P.
PR 04-JUN-1999: 99US-0137502P.
PR 07-JUN-1999: 99US-0137724P.
PR 08-JUN-1999: 99US-0138094P.
PR 10-JUN-1999: 99US-0138540P.
PR 10-JUN-1999: 99US-0138847P.
PR 14-JUN-1999: 99US-0139119P.
PR 16-JUN-1999: 99US-0139452P.
PR 16-JUN-1999: 99US-0139453P.
PR 17-JUN-1999: 99US-0139492P.
PR 18-JUN-1999: 99US-0139454P.
PR 18-JUN-1999: 99US-0139455P.
PR 18-JUN-1999: 99US-0139456P.
PR 18-JUN-1999: 99US-0139457P.
PR 18-JUN-1999: 99US-0139458P.
PR 18-JUN-1999: 99US-0139459P.
PR 18-JUN-1999: 99US-0139460P.
PR 18-JUN-1999: 99US-0139461P.
PR 18-JUN-1999: 99US-0139462P.
PR 18-JUN-1999: 99US-0139463P.
PR 18-JUN-1999: 99US-0139750P.
PR 18-JUN-1999: 99US-0139763P.
PR 21-JUN-1999: 99US-0139817P.
PR 22-JUN-1999: 99US-0139899P.
PR 23-JUN-1999: 99US-0140033P.
PR 23-JUN-1999: 99US-0140354P.
PR 24-JUN-1999: 99US-0140655P.
PR 28-JUN-1999: 99US-0140823P.
PR 29-JUN-1999: 99US-0140991P.
PR 30-JUN-1999: 99US-0141287P.
PR 01-JUL-1999: 99US-0141842P.
PR 01-JUL-1999: 99US-0142154P.
PR 02-JUL-1999: 99US-0142055P.
PR 06-JUL-1999: 99US-014230P.
PR 08-JUL-1999: 99US-0142803P.
PR 09-JUL-1999: 99US-0142920P.
PR 12-JUL-1999: 99US-0142977P.
PR 13-JUL-1999: 99US-0143542P.
PR 14-JUL-1999: 99US-0143644P.

PR 15-JUL-1999: 99US-0144005P.
PR 16-JUL-1999: 99US-0144085P.
PR 16-JUL-1999: 99US-0144086P.
PR 19-JUL-1999: 99US-0144325P.
PR 19-JUL-1999: 99US-0144331P.
PR 19-JUL-1999: 99US-0144332P.
PR 19-JUL-1999: 99US-0144333P.
PR 19-JUL-1999: 99US-0144334P.
PR 19-JUL-1999: 99US-0144335P.
PR 20-JUL-1999: 99US-0144352P.
PR 20-JUL-1999: 99US-0144632P.
PR 20-JUL-1999: 99US-0144884P.
PR 21-JUL-1999: 99US-0144814P.
PR 21-JUL-1999: 99US-0145086P.
PR 21-JUL-1999: 99US-0145088P.
PR 22-JUL-1999: 99US-0145085P.
PR 22-JUL-1999: 99US-0145087P.
PR 22-JUL-1999: 99US-0145089P.
PR 22-JUL-1999: 99US-0145192P.
PR 23-JUL-1999: 99US-0145145P.
PR 23-JUL-1999: 99US-0145218P.
PR 26-JUL-1999: 99US-0145224P.
PR 27-JUL-1999: 99US-0145276P.
PR 27-JUL-1999: 99US-0145913P.
PR 27-JUL-1999: 99US-0145918P.
PR 27-JUL-1999: 99US-0145919P.
PR 28-JUL-1999: 99US-0145951P.
PR 02-AUG-1999: 99US-0146386P.
PR 02-AUG-1999: 99US-0146388P.
PR 02-AUG-1999: 99US-0146389P.
PR 03-AUG-1999: 99US-0147038P.
PR 04-AUG-1999: 99US-0147204P.
PR 04-AUG-1999: 99US-0147302P.
PR 05-AUG-1999: 99US-0147192P.
PR 05-AUG-1999: 99US-0147260P.
PR 06-AUG-1999: 99US-0147303P.
PR 06-AUG-1999: 99US-0147416P.
PR 09-AUG-1999: 99US-0147493P.
PR 09-AUG-1999: 99US-0147935P.
PR 10-AUG-1999: 99US-0148171P.
PR 11-AUG-1999: 99US-0148319P.
PR 12-AUG-1999: 99US-0148341P.
PR 13-AUG-1999: 99US-0148565P.
PR 13-AUG-1999: 99US-0148684P.
PR 16-AUG-1999: 99US-0149368P.
PR 17-AUG-1999: 99US-0149175P.
PR 18-AUG-1999: 99US-0149426P.
PR 20-AUG-1999: 99US-0149722P.
PR 20-AUG-1999: 99US-0149723P.
PR 20-AUG-1999: 99US-0149929P.
PR 23-AUG-1999: 99US-0149902P.
PR 23-AUG-1999: 99US-0149930P.
PR 25-AUG-1999: 99US-0150566P.
PR 26-AUG-1999: 99US-0150884P.
PR 27-AUG-1999: 99US-0151065P.
PR 27-AUG-1999: 99US-0151066P.
PR 27-AUG-1999: 99US-0151080P.
PR 30-AUG-1999: 99US-0151303P.
PR 31-AUG-1999: 99US-0151438P.
PR 01-SEP-1999: 99US-0151930P.
PR 07-SEP-1999: 99US-0152363P.
PR 10-SEP-1999: 99US-0153070P.
PR 13-SEP-1999: 99US-0153758P.
PR 15-SEP-1999: 99US-0154018P.
PR 16-SEP-1999: 99US-0154039P.
PR 20-SEP-1999: 99US-0154779P.
PR 22-SEP-1999: 99US-0155139P.
PR 23-SEP-1999: 99US-0155486P.
PR 24-SEP-1999: 99US-0155659P.
PR 28-SEP-1999: 99US-0156458P.
PR 29-SEP-1999: 99US-0156596P.
PR 04-OCT-1999: 99US-0157117P.
PR 05-OCT-1999: 99US-0157753P.

PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158929P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161935P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 47.2%; Score 1128.5; DB 3; Length 461;
Best Local Similarity 51.6%; Pred. No. 2,2e-85;
Matches 230; Conservative 78; Mismatches 129; Indels 9; Gaps 4;

QY 34 VPAMTREVVDHAKMGREALPMAALTYERAQYLKKAADIIEKDKEIATYLAKEISKA 93
DB 12 VOACTOEYVNAVMEIASKSAQKSWAKPTLWKRAELHKAALIKDNVAPMAESLVEKEIAKP 71
QY 94 YNASTVEVVRTADLIRYAEEGRILSTADEGKM--DASTGH--KLAIVIRQPIGVYL 148
DB 72 AKDSVTEVRSDDLISYCAEEGVRI--LGEGRFLISDSPFGNDRTKYCLTSKPIGLVYL 128
QY 149 AIAPIVYVNLGSGKIPALIGANNVMPKPTQSGVSGVLAKAFAPAGLPAQVNTTNG 208
DB 129 AIPPTNYVNLAVSKIAPIALINGNSLVKRPPTQGANVSCILMHVCHFLAAGFPKGLISCTTG 188
QY 209 RGSRTIGDYIVHEEVNFINTGSTPYGORIGKLGNRPIMLEIGCKDAGIVLADADLDNA 268
DB 189 KSEIIGDFLTMHPAVNCISFTGG-DTGISISKAGMIFPOMELGCKDACIVLDDADLDLV 247
QY 269 AQOYAGADVSGORCTAIKRYLVVEEVADELAEKISENAVAKISVCDPPDNATVTPVID 328
DB 248 ASNIITKGGSYSGORCTAVKVVYLWVESVADBEVKEKAVAKYKLTJVGPPRENSDITAVSE 307
QY 329 NSADFIETSLVVDAROKAKELNEBFKDRGLTLPGLFDHYLTLMKLAWEPPGPIPIIRV 388
DB 308 SSANITEGLVMAKKEGATTCOBYKREGNLIMPFLIDNVRPDKRIAMEBPPGCVFVFLXI 367
QY 389 KDAEEVAVALNKSDPGLQSGSVFTRPQKAFDIANKLEVGTVIHNNKTGRGPNFPFLIGLK 448
DB 368 NSEVEGINHCNASNFGLOCGVFTKDIINKAILISDAMEITGVQNSAPARGPDHPFPGIGLK 427
QY 449 GSGAGVQGIIRYSIEAMTINKSYLDM 474
DB 428 DSGISGQGVTSINIMTKTKVTIVINL 453

RESULT 28
AAG49552
ID AAG49552 standard; protein; 438 AA.

XX XX
AC AAG49552;
XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 62699.
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62699.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134321P.
PR 14-MAY-1999; 99US-0134322P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136329P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.

Qy	261	GORCAIRKRVLYVEEVADEPLAKTISNNVAKLSVSGDFPDNATATPTVYDNDNSADPTSELVYD	340
Db	237	GORCAIRKRVLYVMEESVADBEVBERKAKAKAKLTVGGPEENSDDTAAYVSESSANPIEGELVYD	296
Qy	341	AROKGAKELNEKRRGRRLTTPGLFDHVTLDMLKLAEEPPGPIPLIRVRDAAEAAVIAANK	400
Db	297	AKKKATTCQGEYKREKGNLWPLLDNVRPDMMIAWEEPIGCPVPVYARINSVEGGINHCNA	356
Qy	401	SDFGIQQSVFTFRDFOKAPDIANKLEEVGTVHINNKTGRGPDNPPFLGLKSGSAGAVGGRYS	460
Db	357	SNFGIQQGVFTFDINKKALISDAMETGTVQINSAPARGPDHFFPQGLKDSGIGSGCVTNS	416
Qy	461	IEAMTNVKSIVLDM	474
Db	417	INIMTKVKTYVNL	430
RESULT 29			
AAG32609			
ID	AAG32609	standard; protein; 438 AA.	
XX	AC	AAG32609;	
XX	DT	17-OCT-2000 (first entry)	
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 39373.	
XX	KW	Protein identification; signal transduction pathway; metabolic pathway; hydridisation assay; genetic mapping; gene expression control; promoter; termination sequence.	
XX	OS	Arabidopsis thaliana.	
XX	EN	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	PF	25-FEB-2000; 2000EP-00301439.	
XX	PR	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	16-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	28-APR-1999;	99US-0130891P.	
PR	30-APR-1999;	99US-0131449P.	
PR	04-MAY-1999;	99US-0132048P.	
PR	05-MAY-1999;	99US-0132484P.	
PR	06-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	07-MAY-1999;	99US-0132487P.	
PR	11-MAY-1999;	99US-0132863P.	
PR	14-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	18-MAY-1999;	99US-0134370P.	
PR	19-MAY-1999;	99US-0134768P.	
PR	20-MAY-1999;	99US-0134941P.	
PR	21-MAY-1999;	99US-0135124P.	
PR	24-MAY-1999;	99US-0135353P.	
PR	25-MAY-1999;	99US-0135629P.	
PR	25-MAY-1999;	99US-0136021P.	

PR	7-MAY-1999;	99US-01363832P
PR	28-MAY-1999;	99US-01372722P
PR	01-JUN-1999;	99US-01372222P
PR	03-JUN-1999;	99US-01375282P
PR	04-JUN-1999;	99US-01375022P
PR	07-JUN-1999;	99US-01377242P
PR	08-JUN-1999;	99US-01380094P
PR	10-JUN-1999;	99US-01385400P
PR	10-JUN-1999;	99US-01388477P
PR	14-JUN-1999;	99US-01391119P
PR	16-JUN-1999;	99US-01393423P
PR	16-JUN-1999;	99US-01394532P
PR	17-JUN-1999;	99US-01394922P
PR	18-JUN-1999;	99US-01394642P
PR	18-JUN-1999;	99US-01394616P
PR	18-JUN-1999;	99US-01394366P
PR	18-JUN-1999;	99US-01394577P
PR	18-JUN-1999;	99US-01394588P
PR	18-JUN-1999;	99US-01394599P
PR	21-JUN-1999;	99US-01398177P
PR	22-JUN-1999;	99US-01398039P
PR	23-JUN-1999;	99US-01403533P
PR	23-JUN-1999;	99US-01403544P
PR	24-JUN-1999;	99US-01406355P
PR	28-JUN-1999;	99US-01408233P
PR	29-JUN-1999;	99US-01409511P
PR	30-JUN-1999;	99US-01412877P
PR	01-JUL-1999;	99US-01418422P
PR	01-JUL-1999;	99US-01421544P
PR	02-JUL-1999;	99US-01420555P
PR	06-JUL-1999;	99US-01423300P
PR	08-JUL-1999;	99US-01428033P
PR	12-JUL-1999;	99US-01429270P
PR	12-JUL-1999;	99US-01429777P
PR	13-JUL-1999;	99US-01435642P
PR	14-JUL-1999;	99US-01436524P
PR	15-JUL-1999;	99US-01440055P
PR	16-JUL-1999;	99US-01440085P
PR	16-JUL-1999;	99US-01440066P
PR	19-JUL-1999;	99US-01443335P
PR	19-JUL-1999;	99US-01443332P
PR	19-JUL-1999;	99US-01443311P
PR	19-JUL-1999;	99US-01443332P
PR	19-JUL-1999;	99US-01443333P
PR	21-JUL-1999;	99US-01445086P
PR	21-JUL-1999;	99US-01445088P
PR	21-JUL-1999;	99US-01445087P
PR	22-JUL-1999;	99US-01445087P
PR	22-JUL-1999;	99US-01445089P
PR	22-JUL-1999;	99US-01451512P
PR	23-JUL-1999;	99US-01451515P
PR	23-JUL-1999;	99US-01452145P
PR	23-JUL-1999;	99US-01452248P
PR	26-JUL-1999;	99US-01452767P
PR	27-JUL-1999;	99US-01459133P
PR	27-JUL-1999;	99US-01459188P
PR	27-JUL-1999;	99US-01459191P
PR	28-JUL-1999;	99US-01459611P
PR	02-AUG-1999;	99US-01463386P
PR	02-AUG-1999;	99US-01463388P
PR	02-AUG-1999;	99US-01463389P
PR	03-AUG-1999;	99US-01470387P

PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-01332484P.
PR 05-MAY-1999; 99US-01332485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142877P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148314P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148656P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154729P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160800P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.1%; Score 1102.5; DB 3; Length 438;
Best Local Similarity 50.6%; Pred. No. 3e-83;
Matches 220; Conservative 78; Mismatches 126; Indels 11; Gaps 4;

QY 46 MKAGREALPAMALTYERAQYVHKAADIIENDKEEIAIVTAKEISKANVATEVVRRA 105
1 MDAKVAQAKAMATPLMKRAELHKAALIKHEKAPLAECTIVEIAKPAKDAVSEVVRSG 60
106 DLIRVAEEGIRLSTANDEGKK---DASTG---HKLAVRRQPVGIVLAIAFYNTPVNL 159
61 DLVSYAEAGVRLIGE---GKLLVSDSPFGNERNKYCLSSKILPLGVLAIPFPNYPVNL 116
QY 160 SSGKIAPALIGVNVWFKPPTQGSVGLVYAKAFAGVNTTGRSGRIQYIYE 219
117 AUSKIPALINAGNALVUKPTQCAVAAALHWVHCFHLNAGFPKGLISCVTKGSEIGDPLTM 176
QY 220 HEEVNFINTGSTPVQORIGIKLAGMRPIMLEGGKDAIVLADADLDNNAKOIVAGAYDY 279
177 HPGVNCISFTGG-DTGIAISKKAGWVPLQWELGGKQACIYLEADDDLVSANIVKGFST 235
QY 280 SSGRCIAIKRVLYVEEVADELAKEISENNAKLSVSGPFPNATVPIYIDNSADPISLIV 339
236 SSGRCIAIKRVLYVEEVADELAKEISENNAKLSVSGPFPNATVPIYIDNSADPISLIV 339
QY 340 DAROKAGALENEFKRGRLITPGLFDHYTLDMKLAMEBPPGPIPLTIIRVYDAEAVAIAN 399
296 DAKKGAATTCQETRRKGNLIWPLLDVHRPDMRIAMEBPPGVLPIRINSVEEGIHHCN 355
QY 400 KSDPGLQSSVFTDFQKAFDIANKLEVGTVHINNKGKRGDPNPFGLKSGAGVQIRY 459
356 ASNFVGIQGCIFTRDINKALILISAMETGYQINSARAGDHPFQGLDSDGIGSGITN 415
QY 460 STEAMTNVKSIVIDM 474
DB 416 SINMTKVKXSTVINL 430

RESULT 31

AAG05754 standard; protein; 438 AA.

AAG05754;

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 2277.

KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX XX
OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-00301439.
XX XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.

Db 357 SNFGLOGCVFTKDINKAIIISDAMEITGYQINSAPRGDPHFPGILKDSGISGQVNTS 416
Qy 461 IEAMTNVKSIVLDM 474
Db 417 INLMTKVKTTVINL 430

RESULT 32
AAG05755 standard; protein; 402 AA.
ID AAG05755;
AC AAG05755;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 2278.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2278.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131499P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0137822P.
PR 01-JUN-1999; 99US-0137822P.
PR 03-JUN-1999; 99US-0137828P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147034P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147193P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.

PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150848P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0158293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161929P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.

OY		137	A V I R R Q P G V I A L A M Y N P V U L S S K T A P L I G N V M F P R P O G S V G U L A K A P A E A	156
Db		58	C L T S K S I P L G V L A I P P E N T P V A L A S K A P A L I A N S I V L E K P P O G A V S C L H M H C F H L A	117
OY		197	G L P A G V E N T I T G R G S E I D G Y I V E H E E V A N F I N T S T P V G R I G L A G M R P I M L E G K D A	256
Db		118	G P R K G I S C I T G K S E I G D L T M H P A V N C I S F T G G - D G I S I S K A C M I P L O M E L G G X D A	176
OY		257	G I Y L A A D I D M N A K O J V G A Y D S G O R C T A I K R V I V E E V D E L A E K I S E M V A K L S V G D P	316
Db		177	C I Y L D D D D L B V A S N I T I K G F S Y S Q R C I A K V A V L M E S V A D E L V E R K A K A V A L T I V G P P	236
OY		317	F D N A T P Y I D N S A D F I E S I V D A R O K A E L I N E F K R D G L L T P G L F D H V T L D M K L A W E	376
Db		237	E E N S D I T A V S E S A N F I E G L V M D A K E G A F C Q E Y K R E G A L I W P L L I D V N R P D M R I A W E	296
OY		377	E P R G P L L I I R V M D A E A V A I N K S D P G L O S S V F P R D O R K A F D I A N L L E V G T Y I I N N K T G	436
Db		297	E P R G P V A P V A R I N S V E B G I N H C N A S N F G I Q C V P F K D I N K A I L I S D A M E T G T V O I N S P A	356
OY		437	R G P D N F P F L G L K G S G A V G O I R Y S I E A M T N Y K S I V L D M	474
Db		357	R G P D H P P O G L K O S G I G S G V T N S I N L M T K V T Y I V I N L	394

RESULT 33
 ABU48897
 ID ABU48897 standard; protein; 475 AA.
 AC
 XX ABU48897;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #34424.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 OS Ureaplasma urealyticum.
 XX
 PN W0200277183-A2.
 PD
 XX 03-OCT-2002.
 PF 21-MAR-2002; 2002WC-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 XX 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Hasselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-0299926/02.
 DR N-PSDB; ACA52767.
 XX
 PT New Antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 76821; 1766bp; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation; or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 475 AA:

Query Match 36.3%; Score 868.5; DB 6; Length 475;
Best Local Similarity 39.7%; Pred. No. 1.2e-63;
Matches 190; Conservative 95; Mismatches 177; Indels 17; Gaps 5;

QY 5 YQNVVNGEMKSNVQJELISPIDSSLGFPVPMATRE-EVDHAKKAGREALPMAALTYVE 63
DB 3 YKLLINGAFDAEKKLPVNPSPNNQIAYVNPINHEINTEITPENAHIALKKDPIPKX 62
QY 64 RAQYLKKAADIIRDEKEIATVLAKEISKAYNSVTEVATDLIRYAEGIRLSTSA- 122
DB 63 KCLLLKADKDLDEHKOELAQIISTEIAKGLKSLFEVSAUYLETVEYQKLMQKPI 122
QY 123 --DEGGMDASTGH---KLAIVIRQPVGIVLAIAPNYVNLGSKIAPALIGNVVMP 176
DB 123 IFDE-----TVHVNKNVATYTRIPGVVLAICPNYVNLISLAPALVSGNSLVY 175
QY 177 KPPTQGSVSGVLAKAFAGAGLPAQVFNITTGSGSEIGDYIYHEEYVNFNFGSTIPVQ 236
DB 176 KPSTQSLTIGIRSELVEHGFPGVNNCTTETARITGDLVNNKVKALISFTGGPVGNN 235
QY 237 RICKLAKMRPIMLELGGKAGIYLAADLDLNAKQIYAGAYDSGQRCTAIKKVLVEEV 296
DB 236 HIMEITSKISLVELEGKDALVDDDFELANNEIVKGYGSGQRCTAIKKVFSHKM 295
QY 297 ADRLAEKISENVAKLSVGPDPDNATVTVPIDNSADPIESLVDVDAQKGA--KELNEFKK 354
DB 296 HDLLVNLINKKVDALTVGLPQNPITPLINSNLSIKNNLSLVEDAIKKGALVHAKIVNNE 355
QY 355 DGEILTPGLFDVHTLDMKLAWEPPGPIPIIVKQAEBAVAIAKNSDFGLQSSVFTDRF 414
DB 356 KNNILPLVLDVNTKKEKRVAMEEPPGPIPIITYNSIQEALIDINSQYQLQACIETWY 415
QY 415 QKAFDLANKLEVGTVHNNKTGSGPDPFGLGKSGAGVQIGRISLEAMTNKSVLD 473
DB 416 ASIEQLAQIESGINTINSSSRGPDLTPFGVKGSGFVGQIVDALISMTIKGITIN 474

RESULT 34
AAE37093 standard; protein; 492 AA.
XX

AC AAE37093;

XX 07-AUG-2003 (first entry)

XX Human 9136 ALDH domain consensus protein.

XX

KW Human; aldehyde dehydrogenase; 9136 protein; cardiovascular disorder;
KW cellular proliferative disorder; differentiative disorder; anorectic;
KW endothelial cell disorder; rheumatoid arthritis; arrhythmia; cardiac;
KW heart failure; metabolic disorder; obesity; diabetes; tissue typing;
KW gene therapy; cytosolic; antiinflammatory; cancer; angina; ALDH.
XX Homo sapiens.
XX WO2003033660-A2.
XX
XX 24-APR-2003.
XX
XX 16-OCT-2002; 2002WO-US032971.
XX
XX 16-OCT-2001; 2001US-0329899P.
XX
XX (MILL-) MILLENIUM PHARM INC.
XX
XX Hunter JJ;
XX
XX WPI; 2003-393516/37.
XX
XX Identifying an agent that modulates the level or activity of 9136
XX polypeptide or nucleic acid molecule in a cell, useful for treating
XX cancer, comprises contacting an agent capable of expressing the
XX polypeptide or nucleic acid molecule.

XX Disclosure; Fig 3; 137pp; English.

CC The invention relates to a member of human aldehyde dehydrogenase (ALDH)
CC family, 9136 protein. The invention also relates to a method for
CC identifying an agent that modulates the level or activity of 9136
CC polypeptide or nucleic acid molecule. The methods, polypeptides and
CC nucleic acid molecules are useful for treating a cellular proliferative
CC or differentiative disorder, e.g. cancer; endothelial cell disorder such
CC as rheumatoid arthritis; cardiovascular disorder such as arrhythmia,
CC angina, heart failure; or metabolic disorder such as obesity or diabetes.
CC The methods are also useful in diagnostic assays, predictive medicine and
CC tissue typing. 9136 DNA is used in gene therapy. The present sequence is
CC human 9136 ALDH domain consensus protein

XX Sequence 492 AA:

Query Match 32.9%; Score 786.5; DB 6; Length 492;
Best Local Similarity 38.6%; Pred. No. 8.9e-57;
Matches 193; Conservative 92; Mismatches 166; Indels 49; Gaps 16;

QY 12 EKKSSVN--QIFILSPIDSS--LGPVPMATREVDHAKKAGREAL--PMAALTYVERA 65
DB 1 EWDSSASGKTFEVNPNANGEVYGRPEATADVDAAVPAKAEAFKSGPMMAKVPASERA 60
QY 66 QYLHKRAADIIRDEKEIATVLAKEISK--AYNASVTEVVRTADLIRYA-----EEG 115
DB 61 RILRLKADIIREREDIALTLTDLOKPLAEAKGDTVEGRALDEIRYVAGMARKMGERR 120
QY 116 I--RLSTSDGEGKMDASTGHKLAIVIRQPVGIVLAIAPNYVNLGSKIAPALIGNV 173
DB 121 VIPLSLATDDE-----ELNVTRRRBPGLVGVVISPNNPPLALMLKLAAPLAAQNT 170
QY 174 VNFKPPTQGSVSGVLAKAFAGAG--LPAGVFNITTGSGSEIGDYIYHEEYVNFNFG 230
DB 171 VVLPKPEQPTPLTALLAELEIBEGANNLPRKGVNVVPGFAGVGOALLSHPIDIKLISFG 230
QY 231 STPVGQRIKLA---GMRPIMLELGGKAGIYLAADLDLNAKQIYAGAYDSGQRCTAI 287
DB 231 STEVGLIMEAABAANKLKVITIELGSKSPVYIFDDADLDKAVERIVFGAFGNAQYCTAP 290
QY 288 KRVLVVEEVADELAEKISENVAKLS--VEDPPNAT--VTPVIDNSADPIESLVDVDAQK 344
DB 291 SRLVHESITVDEFEVEKKEKRVKLIIGDPLSDTNIVGPILSQOFDVLSTIEDGKEE 350
QY 345 GAK-----ELNEPKRDGRLLTPGLFDVHTLDMKLAWEPPGPIPIIRVYDAEBAVA 396


```

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151308P.
PR 07-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 13-SEP-1999; 99US-0153707P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158322P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

```

Query Match 32.8%; Score 784.5, DB 3, Length 293;
 Best Local Similarity 52.4%; Pred. No. 6,3e-57;
 Matches 150; Conservative 51; Mismatches 84; Indels 1; Gaps 1;

QY 189 LAKAPAEAGLPACVPTTITRGSEIGDIYVEHEBVNFINTGSTPTPGORIGKLAGRPTM 248

```

Db 1 MVHCFHLAEPKGLISCVTGGKSEIGDFLTMHPGVNCISFTGG-DTGIAISKAGVNPQ 59
QY 249 LEIGKDAQIVLADADLDNNAKOIYAGAYDYGQRTAIKRVLYVEVDELAEKISENV 308
Db 60 MEIGKDAQIVLEADLDLVSAIVYGFSGYSGQRTAVKVVLYMESIDAVQKNAVL 119
QY 309 AKLSVDPFPDNAVTPVIDNSADPFESIYVVAROKAGELMEFKRDGLTPGPEPDHT 368
Db 120 AKLVGPPEDSDITPVVTESSNFTLGLVMDAKEGATPCQRYRREGNLLPDLIDHYR 179
QY 369 LPMKLAMEBPFGLPIIIRKDAEBAVALANKSDPGLQSSVFTRPFOKAFDIANKLEVGT 428
Db 180 PMRIAMEBPFGLPVLRINSVEBGHHCNNSNFGLOCCITFRDINKAILLSDAMEGT 239
QY 429 VHINKTGRGPDNFPPLGLKSGAGYQGIIRYIEAMTVKSYVLDM 474
Db 240 VOINSAPARGPDHPFPGGLKDSGIGSQGITNSINMTTKVKSIVINL 285

RESULT 36
AAE39889
ID AAE39889 standard; protein; 493 AA.
XX
AC AAE39889;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human aldehyde dehydrogenase domain consensus peptide.
XX
KW Human cytochrome p450; dehydrogenase/reductase; osteomalacia; psoriasis;
KW lipoxigenase; hydratase; proliferative disorder; hematopoietic disorder;
KW differentiative disorder; carcinoma; sarcoma; leukemia; immune disorder;
KW anorexia nervosa; neural degeneration; muscular disorder; lipid disorder;
KW multiple sclerosis; encephalomyelitis; metabolic disorder; osteoporosis;
KW myasthenia gravis; bone metabolism; immunomodulator; obesity; anorectic;
KW eating disorder; osteodystrophy; arthritis; diabetes; anabolic; rickets;
KW milk fever; diabetes mellitus.
XX
OS Homo sapiens.
XX
PN US2003092658-A1.
XX
PD 15-MAY-2003.
XX
PF 20-JUN-2002; 2002US-00175696.
XX
PR 02-FEB-2001; 2001US-0266140P.
XX
PR 04-FEB-2002; 2002US-00067668.
XX
PA (MEYER/) MEYERS R E.
PA (GLUC/) GLUCKSMANN M A.
PA (RUDO/) RUDOLPH-OWEN L A.
XX
PI Meyers RE, Glucksmann MA, Rudolph-Owen LA;
XX
DR WPI; 2003-765490/72.
XX
XX
XX New cytochrome P450 family members, dehydrogenase/reductase protein,
PT lipoxigenase family member and human hydratase useful for treating
PT cellular proliferative and/or differentiative disorders.
XX
XX Disclosure; Fig 16; 0pp; English.
XX
XX The invention relates to isolated cytochrome p450 family members such as
CC 33312, 33302 and 32579, dehydrogenase/reductase protein such as 21509,
CC 33770, lipoxigenase family member such as 46638 and human hydratase such
CC as 50090. The invention is useful as diagnostic targets and agents for
CC controlling one or more of cellular proliferative and/or differentiative
CC disorders such as carcinoma, sarcoma or haematopoietic disorders such as
CC leukemia, immune disorders such as diabetes mellitus, arthritis,
CC multiple sclerosis, encephalomyelitis, myasthenia gravis, psoriasis, etc.
CC or metabolic disorders such as obesity, anorexia nervosa, cachexia, lipid

```

disorders and diabetes. The invention is useful for controlling disorders associated with abnormal fatty acid biosynthesis and neural degeneration. The invention is useful for controlling disorders associated with defects in fatty acid oxidation or proliferation or muscular disorders and for treating disorders associated with bone metabolism such as osteoporosis, osteodystrophy, osteomalacia, rickets, osteitis fibrosa cystica, milk fever. The present sequence is human aldehyde dehydrogenase domain consensus peptide

SQ Sequence 493 AA;

Query Match	32.5%	Score 777	DB 7	length 493
Best Local Similarity	38.5%	Pred. No. 5.5e-56		
Matches 193; Conservative	92	Mismatches 166	Indels 50	Gaps 17

```

Qy 12 MMKSSVN--QIILSPIDSS--LGFPVPMTRSEVDHMAKGRAL---PAMALTYTERA 65
Db 1 EVDSDASGKTFEVDNPNANGEVIGRPEPEATAEVDDAVNAKAKAFKSGPMWAVPASERA 60
Qy 66 QYLHSAADIIEERDKESIAIVLAKESIK--AYNASVTEVVRTADLIRYAA-----EEG 115
Db 61 RILRKADLIEERDEBLALETLDLCKPLAEAKGDTVEGGAIDEIITYYAGMARKLMGER 120
Qy 116 I--RLSTADEGGRNDASTGHKGLAVRRQPVGIVLAIAPNRYVNISSGSLAALLIGNV 173
Db 121 VIPSATGDE-----ELMYTRREPLGVGVGISPNWFELLALAMGLAALAAAGNT 170
Qy 174 VMFKPPTGSSVAGVLAKAFEAAG--LPAQVFNITGRGSIBIDYIVHEHSEVNFINTPT 230
Db 171 VVKRSEGPPLTALLAELIEEGANNLPKGVNVNVPFGGAEVGOALLSHPIDIKISFTG 230
Qy 231 STPVQORIGKLA---GMRPIMELGKGDAGIYLAADDLDNNAKOIYAGAYDSGQRCTAI 287
Db 231 STEVKLIMEAANAANKLKKVLTLELGKSSPVIAFDDLDLKAVERIYVFGAFGNAQVCIA 290
Qy 288 KRVLVVEAEDEBLAEKISENVAKLS--VGDPFDNAT--VTPVIDNNGAD--FISELVVDARQ 343
Db 291 SRLVHESIYDFVEVLEKRVKKLKIQPLDSDTNIYGPILISEQGFDRVLMYIIEDK 350
Qy 344 KGAK-----ELNEFKRQGRLLTGLFDHVTLDNKLMAEFPFGTILPIIRYKDAEAV 395
Db 351 EGAKVLGGERDESKELYLGAGYVVOPIITFDVTPDMKIMKEEITFGVLPPIRKEDIDEAL 410
Qy 396 AIANKSDGLOSSVTRD--FOKAFDIANKLEVGTVHNNKTKGRGPD--NPFFELTK--GSQA 452
Db 411 ELANTVEGLAYVTFKDLIARAFRAKALIEAGIYVWNDVCHAAARPQLPFGGVKQSSGI 470
Qy 453 GVQ--GIRYSIEAMTVKSLIVL 472
Db 471 GREHGKYGLEEVTEIKIVTI 491

```

RESULT 37
ADQ48417
ID ADQ48417 standard; protein; 493 AA.

AC ADQ48417;

DT 23-SEP-2004 (first entry)

Aldehyde dehydrogenase domain consensus, SEO ID 21.

KM Neuroprotective; Nootropic; Cerebroprotective; Anticonvulsant;
KM Antiparkinsonian; Hepatotropic; Antiinflammatory; Osteoprotic;
KM Antipneumatic; Antiarthritic; Cytostatic; Antidiabetic; Antithyroid;
KM Gastrointestinal; Antulcer; Cardiac; Immunosuppressive; Dermatological;
KM Virucide; Hemostatic; Gene therapy; human; cytochrome P450;
KM cytochrome P450 3312; cytochrome P450 3303; cytochrome P450 32579;
KM dehydrogenase; reductase; dehydrogenase/ reductase 21509;
KM dehydrogenase/ reductase 33770; lipoxxygenase; lipoxxygenase 46338;
XX hydrolase; hydrolase 50090.
XX Unidentified.
XX

XX	US2004132087-A1.
PN	
XX	
PD	08-JUL-2004.

PF 11-FEB-2004; 2004US-00776871.

PR	13-MAR-2002	2000US-01939260P
PR	13-MAR-2002	2000US-0206675P
PR	19-MAY-2000	2000US-0125567P
PR	15-JUN-2000	2000US-0211777P
PR	02-FEB-2001	2001US-0266140P
PR	30-MAR-2001	2001US-0062330P1
PR	02-APR-2001	2001WO-US010720
PR	21-MAY-2001	2001US-00862658
PR	21-MAY-2001	2001WO-US016380
PR	15-JUN-2001	2001US-00882837
PR	15-JUN-2001	2001WO-US019319
PR	04-FEB-2002	2002US-00067668
PR	20-JUN-2002	2002US-00175696

PA (MILL-) MILLENNIUM PHARM INC.

PI Meyers RE, Glucksmann MA, Rudolph-Owen LA;

DR WPI; 2004-533235/51.

PT New 33312, 33303, 32579, 21509, 33770, 46638, or 50090 nucleic acid
PT molecules, useful for treating genetic disorders of the membrane
PT transport, CNS disorders, liver disorders, skeletal muscle disorders, or
PT cancer.

PS Disclosure; SEQ ID NO 21; 248bp; English.

CC The present invention relates to novel human coding sequences and
CC proteins for cytochrome P450 proteins 33312 (ADQ48397-ADQ48399), 33303
CC (ADQ48400-ADQ48402) and 32579 (ADQ48403-ADQ48405); dehydrogenases/
CC reductases 21509 (ADQ48409-ADQ48411) and 33770 (ADQ48412-ADQ48414);
CC lipoygenase 46638 (ADQ48418-ADQ48420) and hydrtase 50090 (ADQ48424-
CC ADQ48426). The sequences are useful for diagnosing and treating
CC disorders, such as genetic disorders of the membrane transport
CC (aminocidurias, cystinosis), CNS disorders (Alzheimer's disease,
CC epilepsy, Parkinson's disease), liver disorders, skeletal muscle
CC disorders, cellular proliferative and/or differentiative disorders
CC (cancer), hormonal disorders (diabetes, thyroid disorders), immune and
CC inflammatory disorders (rheumatoid arthritis, osteoarthritis, ulcer),
CC cardiovascular disorders, blood vessel disorders, neutrophil disorders
CC (neutropenia, lupus), testicular disorder (mumps) and platelet disorders
CC. The present sequence is a dehydrogenase consensus sequence used in a
CC sequence alignment to illustrate the invention.

Sequence 493 AA;

Query Match	32.5%	Score 777;	DB 8;	Length 493;
Best Local Similarity	38.5%;	Pred. No. 5.7e-56;		
Matches 193; Conservative	92;	Mismatches 166;	Indels 50;	Gaps 17

Qy	12	EMKSSVW--QIETLSPIDDS--LGVVPAMTREVUDHAKKAREL---PAMALLTYVERA	65
Db	1	EWWDSDASGKTFEEVWPANKGEVIGRVPATIEDVDAAVKAKEAFKSGPMMAKVPASERA	60
Qy	66	QYLHKADIIERKDEIATVLAKESK--AANASTEVTPTADILRYAA-----EEG	115
Db	61	RIIRKADLIBEREEDIALETLDDKPLAEKGTVEGRAIDELIRYAGARKLTMERR	120
Qy	116	I--RLSTADEGCKMDASTGHTLAVIRROPVGIVLAIAPVYPVNLGSKTAPALIGNV	173
Db	121	VISLTLATDGB-----ELANTREPELGVGVISPMNFPLLTALMTKLAPALAAAGNT	170
Qy	174	VMKRPYQSGVSLYAKAFAG---LPAGFNTITRGSEIGDIYVEHEVNFIAFTG	230
Db	171	VVLKPEQPPPLTLLAEILIEBAGANNLPKGVVNVVPEFGAEVQALSHPDIDKISFTG	230
Qy	231	STPEVGRIKGLA---GMRPIMLELGCKAGIVLADADIDNNAKOIVAGAYDVGSGRCIAI	287

OS Bacteria.
 XX US2003233675-A1.
 XX
 XX 18-DEC-2003.
 PD
 XX 20-FEB-2003; 2003US-00369493.
 PF
 XX 21-FEB-2002; 2002US-0360039P.
 PR
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS.
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PS Claim 1; SEQ ID NO 1073; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX

SQ Sequence 463 AA;

Query Match 31.7%; Score 759.5; DB 8; Length 463;
 Best Local Similarity 37.1%; Pred. No. 1.5e-54;
 Matches 175; Conservative 95; Mismatches 183; Indels 19; Gaps 8;

QY 8 YNNGEWSKSVNO--IRILSPIDSSLGFPAMTRREVDHAKMAGRALPAMALTYREA 65
 DB 2 FIDGKM--INREMDVINPYSLEVIYKKIPALSRBAKAIIDTAEKYKVMKNLPITTKY 58
 QY 66 QYLHRAADIIERDKIEIATVLAKEISKAYNAVSTEVRTADILRYAAEGIRLSTADG 125
 DB 59 NILMNIAKIKKEKEELAKIIMADGPKIKQARVEVERSIGPKLAA---FVYKERDVE 115
 QY 126 GKMDASTGKLAIVIRROPYGIYLAIPVYPVNLSSKIAIPALIGNVVMFKPPPGSVS 185
 DB 116 IPSD---DRLIFTRREPVGIVCAIRPFNPPLNSHKIAPALATGNVIVHHPSSKAPLV 171
 QY 186 GLVIAK---AFPAEAGLPAGVPTNTINGRSGEIGDYVHEBEVNFINTSTPVGQRIGLU 241
 DB 172 CIELAKIENALKKNVPLGVNLLTGAGEVGVDEIVNEKVMNISTGSSKVGELITTK 231

QY 242 AGMRPIMLEGGKADAGIVLADADIDMNAKOIVAGAYDSGQRCTAIKRVLVEEVADELA 301
 DB 232 AGPKIALIEGGVNPVNIYLVKQADLNKAVNALIKSPFIYAQVCSVGMILVDESIADEFI 291
 QY 302 EKISENVAKLSVGDPEFNAT-VTPYIDNSADFIKLSLVYDARQKAKELNFKKDGRLTT 360
 DB 292 EMFVNKAKVLNVGNPLDEKTVGPLISVEHAEWEKVEKAIIDEGKLLLGKDKKALFY 351
 QY 361 PGLPHTVTLDMKLMABEPPGPILPPIRYKDAEBAVALINKSDPGLQSVFTTRDPQKADI 420
 DB 352 PTLIE-VDRNLIKTEFTFAVPIPIRTNE-BEMIDINSTERGLHSAIFINDINKSLKF 409
 QY 421 ANKLEVGTVHINNKTRGPDNFPPLGLKSGAGVQGIKRSIEMNTVKSIVL 472
 DB 410 AENLEFGGVVINDSSLFQDNNMPCFQVKKSGIGREGVXYAMEMSNIKTIIL 461

RESULT 40

ADN18535 standard; protein; 455 AA.

ADN18535;

02-DEC-2004 (first entry)

Bacterial polypeptide #1188.

Recombinant DNA construct; transformed plant; improved plant property;
 cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 pathogen tolerance; pest tolerance; plant disease resistance;
 cell cycle pathway modification; plant growth regulator;
 homologous recombination; seed oil yield; protein yield; carbohydrate;
 nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 bacterial polypeptide.

Bacteria.

US2003233675-A1.

18-DEC-2003.

20-FEB-2003; 2003US-00369493.

21-FEB-2002; 2002US-0360039P.

(CAOY/) CAO Y.
 (HINK/) HINKLE G J.
 (SLAT/) SLATER S C.
 (CHEN/) CHEN X.
 (GOLD/) GOLDMAN B S.

Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

WPI; 2004-061375/06.

New recombinant DNA construct comprising a promoter positioned to provide
 for expression of a polynucleotide encoding a polypeptide from a
 microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 1188; 122pp; English.

The invention relates to a recombinant DNA construct comprising a
 promoter functional in a plant cell, where the promoter is positioned to
 provide for expression of a polynucleotide encoding a polypeptide from a
 microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,

[illegible][illegible]

US2003233675-A1.

18-DEC-2003.

20-FEB-2003; 2003US-00369493.

21-FEB-2002; 2002US-0360039P.

(CAOY/) CAO Y.
(HINK/) HINKLE G J.
(SLAT/) SLATER S C.
(CHEN/) CHEN X.
(GOLD/) GOLDMAN B S.

Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
WPI; 2004-061375/06.

New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 17; 122bp; English.

The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactanmanan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

```

Query Match      30.8%; Score 737; DB 8; Length 476;
Best Local Similarity 35.0%; Pred. No. 1.2e-52;
Matches 168; Conservative 96; Mismatches 202; Indels 16; Gaps 6;

QY      EYQNVVNEKWSVQVQIEILSPIDSSIGFVPAMTRREVDHMKGREALPAMALTYE 63
Db      3 EKMILIGSEWDEKKEIKIVIYPTRKPIGRVPKDEKQVEXKIEAKGFEKITSIAYE 62

QY      64 RAQYLHKADIIERPEKEIATLAKESIKAVYASTEVVTRADILRYAAEEGIRLS---- 119
Db      63 RYHILMRAAQLLKEBAEFKATLVLEVKTRIATEVQRAQTLIFSAEAKKNGSTF 122

QY      120 -TGADEGSKMDASTGCHKLAIVTRBPQVGIVALAIPNYVPUNTSGSIAPALIGNVVMEKP 178
Db      123 PIDAPNPK-----GKLGFYIRVQVGIVSAITPFPPLNTSMHKVAPALAGNAVILKP 176

QY      179 PTQGSVSGVLAKAPAAEAGLPAGVFNITGSGSEIGDYIVHEEKNFNFPGSTPVGORI 238
Db      177 SERTLTPLTMEGILLENKGVPPKALSVPGIG-DVGKAMTHTPDRAVVSFGSKRVGDI 235

QY      239 GKLAMGRPIMTELGGKDAIGIVLADADLNNAKOIVAGAYDYSGQRCFAIKRYLVVEEVD 298

```

Db	226	ARQVSIKQLVLELGSNSAIIILHKQNLBEAVKVTQGGALIAQVCISIQVRVVEHDDLD	235
Qy	299	ELAEKISENVAKL SVGDEPFDNAT -VTPVIDNSADFIESLVYDARQKAK -ELINEK -R	354
Db	296	EFIRKLKERVSLKVKQDPMNEPDTD GPMIAPSELBERIQWISIEAVQKGIKIEAGGLRCAB	355
Qy	335	DGRLLTPEGLFDHYTLIDMKLAMEEPFGPLIPITRVDAEAAVAIANKSDGLQSSVTRDF	414
Db	356	DETFISPTIVSLVPEDSKLFEKEAPAPVVAWVPYKDIEBAIIMVWSSDGLQGVETNDV	415
Qy	415	QKAFDIAKLEVGTHINNKGRGPDNPFGLKSGSGAGVOGIRISIEAMTVKSIVLDM	474
Db	416	KLWAKCIKEIEAGVVLINBGPTRADHMPYGSVKISGIGRSGPKAIBDYTEIKTIVFDL	475
RESULT 44			
ADS28216			
ID	ADS28216	standard; protein; 482 AA.	
XX	AC	ADS28216;	
XX	DT	02-DEC-2004 (first entry)	
DE	XX	Bacterial polypeptide #17249.	
XX	XX	Recombinant DNA construct; transformed plant; improved plant property;	
KW	KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;	
KW	KW	pathogen tolerance; pest tolerance; plant disease resistance;	
KW	KW	cell cycle pathway modification; plant growth regulator;	
KW	KW	homologous recombination; seed oil yield; protein yield; carbohydrate;	
KW	KW	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;	
XX	OS	bacterial polypeptide.	
XX	OS	Bacteria.	
XX	PN	US2003233675-A1.	
XX	PD	18-DEC-2003.	
XX	PE	20-FEB-2003; 2003US-00369493.	
XX	PR	21-FEB-2002; 2002US-0360039P.	
PA	PA	(CAOY/) CAO Y.	
PA	PA	(HINK/) HINKLE G J.	
PA	PA	(SLAT/) SLATER S C.	
PA	PA	(CHEN/) CHEN X.	
PA	PA	(GOLD/) GOLDMAN B S.	
PI	PI	Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;	
XX	XX	WPI; 2004-061375/06.	
PT	PT	New recombinant DNA construct comprising a promoter positioned to provide	
PT	PT	for expression of a polynucleotide encoding a polypeptide from a	
PT	PT	microbial source, useful for producing plants with improved properties.	
XX	XX	Claim 1; SEQ ID NO 17249; 122bp; English.	
CC	CC	The invention relates to a recombinant DNA construct comprising a	
CC	CC	promoter functional in a plant cell, where the promoter is positioned to	
CC	CC	provide for expression of a polynucleotide encoding a polypeptide from a	
CC	CC	microbial source. The invention also relates to a transformed plant	
CC	CC	comprising the recombinant DNA construct and a method of producing a	
CC	CC	transformed plant having an improved property. The plant is a crop plant	
CC	CC	such as maize or soybean. The method of producing a transformed plant	
CC	CC	having an improved property comprises transforming a plant with the	
CC	CC	recombinant DNA construct and growing the transformed plant, where the	
CC	CC	polynucleotide or polypeptide is useful for improving plant properties.	
CC	CC	The recombinant DNA construct is useful for producing plants with	
CC	CC	improved plant properties, e.g. improved cold, heat or drought tolerance,	
CC	CC	increased resistance to herbicides, extreme osmotic conditions, pathogens or pests,	
CC	CC	increased resistance to plant disease, better growth rate by modification	

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
OS
XX US6551795-B1.
PN 22-APR-2003.
XX
PD 18-FEB-1999; 99US-00252991.
PF 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX MPI; 2003-615309/58.
XX N-PSDB; ABD12127.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 27302; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences AB067826-
XX AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 526 AA;
SQ
Query Match 30.7%; Score 733.5; DB 7; Length 526;
Best Local Similarity 36.8%; Pred. No. 2,6e-52;
Matches 172; Conservative 87; Mismatches 186; Indels 23; Gaps 7;
QY 11 GEMK--SSVNOIILSPIDSSIGFVPAMTREEDVHMKAGRALPAMALTYERAQYL 68
DB 53 GEMRHGRAGRRLKYNPFDGSLLEIEQADRDLDAAVAKAAVOPAMALGPSABAAYL 112
QY 69 HKAAADIIRKXERITATYLAKEISKAYNASTEVRADLRVAEGLRSTGA---DB 124
DB 113 YKAVEFDRRHEHIVMIIRSSG-----TRKAELEWGAARATILTESASFPARVH 163
QY 125 GGMNDSTGHLAVIRROPVGIYLAIPYVNVNLGSGSKIALPILGNNVMPFPTQGSV 184
DB 164 GRIVEDVPKESRVSRAIGVGVISPMFPLHTRQSLAPALAGNAVVPADDTY 223
QY 185 -SGVLAKAPAEAGLPAGVNNTTGRGSEIGDIYVEHEEVNFTFGSTPVGRIKLAG 243
DB 224 CGDILARIFEEAGLPAGLFSVVGSGSEIGDAFVEHPVGLVTFGSGTFVGNIRIAS 283
QY 244 ---MPEIMELGKXAGIYLAADADLDNAKQIVAGAYDVGSGRCATKRVLVESVABE 299
DB 284 GGAHLKVALLELGENSPFVILGDADLEQAVNAVFGKFLHOGQICMAINCITVEDSLYDA 343
QY 300 LAEKISENVAKLSVGGP--FDNAVTPVIDDNSADFLISLVVDAROKGAKELNFKRDR 357
DB 344 FAARFVERKGLVGDPRADTA-VGPIVNARQLEGLERIKRLARQEGAKPLVEGGVDG 402

QY 358 LITPGLFDHVTLDMLKAMEPFGPILPIIRVDAEBAVIANKSPGLOSVFTDFOKA 417
DB 403 LIAHPVGEVYATMETARDEIFGLPGLRADAHLPLANASEGLSSAFPSRDLERA 462
QY 418 PDIANKEVGTVINNNKTGRGPDNFPFLGLKSGAGVOCIRYSIEMNT 465
DB 463 VRFARGLRAGMTHVNDIPVNDKANPFGSEKNSGIGRFNGDAIEEFT 510
RESULT 47
ADSS27421
ID ADSS27421 standard; protein; 470 AA.
XX
XX ADSS27421;
AC
XX 02-DEC-2004 (first entry)
XX
XX Bacterial polypeptide #16454.
DE
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; omomais;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
KW
XX Bacteria.
OS
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
PD
XX
XX 20-FEB-2003; 2003US-00369493.
PF
XX
XX 21-FEB-2002; 2002US-0360039P.
PR
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX MPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1, SEQ ID NO 16454; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan

ID	ABU28726	standard; protein; 482 AA.
XX	ABU28726;	
AC	ABU28726;	
XX	19-UUN-2003	(first entry)
DT	19-UUN-2003	(first entry)
DB	Protein encoded by Prokaryotic essential gene #44253.	
XX	Antisense; prokaryotic essential gene; cell proliferation; drug design.	
KW	Antisense; prokaryotic essential gene; cell proliferation; drug design.	
XX	Escherichia coli.	
OS	Escherichia coli.	
FN	W020027183-A2.	
XX	W020027183-A2.	
PD	03-OCT-2002.	
XX	03-OCT-2002.	
FE	21-MAR-2002; 2002MO-US009107.	
XX	21-MAR-2002; 2002MO-US009107.	
PR	21-MAR-2001; 2001US-00815242.	
PR	06-SEP-2001; 2001US-00948993.	
PR	25-OCT-2001; 2001US-0342923P.	
PR	08-FEB-2002; 2002US-00072851.	
PR	06-MAR-2002; 2002US-0362699P.	
XX	06-MAR-2002; 2002US-0362699P.	
PA	(EIT-) ELITRA PHARM INC.	
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;	
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreysch RA, Xu HH;	
XX	WPI; 2003-029926/02.	
DR	N-PSDB; ACN32596.	
XX	N-PSDB; ACN32596.	
PT	New antisense nucleic acids, useful for identifying proteins or screening	
PT	for homologous nucleic acids required for cellular proliferation to	
PT	isolate candidate molecules for rational drug discovery programs.	
XX	isolate candidate molecules for rational drug discovery programs.	
PS	Claim 25; SEQ ID NO 56650; 1766bp; English.	
XX	Claim 25; SEQ ID NO 56650; 1766bp; English.	
XX	The invention relates to an isolated nucleic acid comprising any one of	
CC	the 6213 antisense sequences given in the specification where expression	
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:	
CC	(1) a vector comprising a promoter operably linked to the nucleic acid	
CC	encoding a polypeptide whose expression is inhibited by the antisense	
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated	
CC	polypeptide or its fragment whose expression is inhibited by the	
CC	antisense nucleic acid; (4) an antibody capable of specifically binding	
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular	
CC	proliferation or the activity of a gene in an operon required for	
CC	proliferation; (7) identifying a compound that influences the activity of	
CC	the gene product or that has an activity against a biological pathway	
CC	required for proliferation, or that inhibits cellular proliferation; (8)	
CC	identifying a gene required for cellular proliferation or the biological	
CC	pathway in which a proliferation-required gene or its gene product lies	
CC	or a gene on which the test compound that inhibits proliferation of an	
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a	
CC	compound's activity; (11) a culture comprising strains in which the gene	
CC	product is overexpressed or underexpressed; (12) determining the extent	
CC	to which each of the strains is present in a culture or collection of	
CC	strains; or (13) identifying the target of a compound that inhibits the	
CC	proliferation of an organism. The antisense nucleic acids are useful for	
CC	identifying proteins or screening for homologous nucleic acids required	
CC	for cellular proliferation to isolate candidate molecules for rational	
CC	drug discovery programs, or for screening homologous nucleic acids	
CC	required for proliferation in cells other than S. aureus, S. typhimurium,	
CC	K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of	
CC	the target prokaryotic essential genes. Note: The sequence data for this	
CC	patent did not form part of the printed specification, but was obtained	
CC	in electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences	
XX	ftp.wipo.int/pub/published_pct_sequences	
XX	Sequence 482 AA;	

Best Local Similarity 34.3%, Pred. No. 9.8e-52; Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;

```

QY 6 QNYVNGEKKSSVN--QIEILSPIDDSIGCFPMPTREEVHAMKAGREALPAMALTYVE 63
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 12 QALINGEVLDA NNGEALIDVTNPNANGDLGSPFKGADETRAITDAARALPAMRALTAKE 71
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 64 RAOYLHKAADIIEEDKEIATVLAKEISKAYNAVSTEVTATDLIRYAAEGIRLSTSD 123
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 72 RATILRMVFNIMMEHODDLARLMTLEOQKPLAEAKGEISYVASFIEWFAEGRKIYDGI 131
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 124 EGGKMDASTGKKLAVITRRQPIGVILALAPHYYPNNLSGSKTAPLIGANNVMPFPQGS 183
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 132 PGHQAD-----KRLIVIKQPIGVTAATTPMNFPAAMITRRKGPALAACTMVLKPAQTP 186
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 184 VSGVLAKAFABAGLPAGFENTITGRGSEIGDYVEHEEVNFIFGTSTPGORIGTLAG 243
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 187 FSALALAEALIRAGVPPGVFNVTGSAAGVNELTSPNPLVRKISFTGSTIEIGRLMEQCA 246
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 244 --MRPIMLELGKQAGIVLADADLDNNAKOIVAGAYVSGQRCAIRVLVEEVADELA 301
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 247 KDIKKVSLELGQGNAPFIVPDDADIDKXAVEGLASKFPNAGQTCACARLIVQDSVYRFA 306
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 302 EKISEBNNAKLSVGPFPDN-ATVTFVIDNSADFTIESLVDARQKAYEL--NEFKDGR 357
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 307 EKLQAAVSKLTIIGDLNGVTIGLILIEKAVAKVEEHIADALEKAGAVVCGKHAHERGN 366
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 358 LLTGLGPLHYVLDLMDKLAEEBPFGPILILIRYKDAEAAVAITANKSDPGLQSSVFFRDRQKA 417
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 367 FFOPIILVDVPANNAKVSKEETFGPLAPLFRPKDADADYIAQANDTEFGALAVFYARDLSRV 426
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 418 FDIANKLEVGTVAHINNTKG-RGPNPFPPLGLKSGSAGVQGIKRSIEAMTNVKSIVLDM 474
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 427 FRVEALEEYGVGIN--TGIISNVAVAFPGIKASGLRBSGSKYGEIDYLEIKYWCICL 482
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 50
ADS45081
ID ADS45081 standard; protein; 482 AA.
XX
AC ADS45081;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #23511.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
WP1; 2004-061375/06.

```

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX
PS Claim 1; SEQ ID NO 23511; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 482 AA;

Query Match 30.3%; Score 726; DB 8; Length 482;
Best Local Similarity 34.3%; Pred. No. 9, 8e-52;

Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;

6 QNVVNGEMSSVNV-QIETILSPIDDSLGFPVPMPTREEDVHMKAREALPMAALTYVE 63
12 QALINGEMLDANNGEAIDVTNPANGKLSVPMGADETRAALDAANRLPAMRALTAKE 71
64 RAQYLHKAADIIERDEKEIATVLAKEISKAYNAVTEVVRTADLIRYAAEGRIETSD 123
72 RATILRNWNLMMHEODDLARLMTLEOGKPLAEAKESIVYASFIEMFAEGRIYGDRT 131
124 EGGKNDASGHTKLAIVRRQPVGIVLAIAFYNPVNLSSGIAPALIGVNVWFKPTQGS 183
132 PEGHQA-----KRLIVIKQPIGVTAITPWNFPAAITRKAGPALAAGCTWVLRKPSQTP 186
184 VSGLVLAKEAFEAAGLPAGVFNTITGSGEIGDIYVHEEYVNFNPGSTPVGORIGKLAG 243
187 FBALALAEALIRAGVAGVFNVTGSAVGNELTNPVKLSFGSTIEIGQLMEQCA 246
244 --MRPIMELGKGDAGIVLADADLDNAAKQIVAGAYDSGORCTAIKRLVVEEVADLA 301
247 KOIKKVSLELGGNAPFTVDDADLDKAVGALASKRRNNGQTCVCANRLVYDGDVDRFA 306
302 EKISNVAAKLSVGDPPDN-ATTPVYIDNSADFISSLVVDAROKAKEL---NEFRDGR 357
307 EKLQCAVSKLHIGDGDNGVTIGPLIDERAVAKEEHIDALEKAGARVCGKAKHRCGN 366
358 LITPGFHDVTLDMKLAWEPEPGILPIIRVYDAEBAVAIANKSDGLOSVFTPFOKA 417
367 FPGPTLIVDPAAKAKSKETFGPLAFLRFKDEADVIAOANDTEGGLAAVFAARLSRV 426
418 FDIANLLEVGTVHINKTG-RGPDNPPLGLKSGAGVQIRYSIEMTNVKSIVLDM 474
427 FAVGEALLEGIVGIN--TGISNEVAPFGIKASGIGRGSSKXGIDYLEIKTMCIQL 482

RESULT 51
ADN46591

ID ADN46591 standard; protein; 511 AA.

XX
AC ADN46591;

XX
DT 01-JUL-2004 (first entry)

XX
DE Thermococcus kodakarensis KOD1 protein sequence Segid469.

XX
KW gene disruption; gene targeting; marker gene; transformation;
KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
KW structure; gene function; enzyme activity; medicine;
KW forensic science; food; drug inspection; molecular biology; immunology.

XX
OS Thermococcus kodakarensis.

XX
PN MO2004022736-A1.

XX
PD 18-MAR-2004.

XX
PF 29-AUG-2003; 2003MO-IB003597.

XX
PR 30-AUG-2002; 2002JP-00319011.

XX
PI (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX
PI Imanaka T, Atomi H;

XX
PI WPI; 2004-257583/24.

XX
PT Method for disrupting targeted gene in genome of organism particularly
PT thermostable bacterium and with genome chips for analysis, applicable in
PT studying gene structure and functions.

XX
PS Claim 9; SEQ ID NO 469; 598bp; Japanese.

XX
CC This invention relates to a novel method for targeting disruption of an
CC arbitrary gene in a genome of an organism which comprises providing the
CC whole sequential data of the genome of such organism, selecting at least
CC 1 arbitrary region in the sequence, providing a vector that contains a
CC sequence homologous with the selected region and a marker gene,
CC transformation, and homologous recombination. The genome is preferably
CC the genome of a hyperthermostable archaeobacterium, particularly
CC Thermococcus kodakarensis KOD1. The method is for targeting the
CC disruption of a gene in the genome of an organism, which is applicable in
CC studying gene structure and functions as well as enzyme activities of
CC encoded proteins and useful in medicine, forensic science, food or drug
CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of a protein
CC encoded by the genome of Thermococcus kodakarensis which was derived
CC using the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.int/pub/published_pct_sequences

XX
SQ Sequence 511 AA;

Query Match 30.2%; Score 722; DB 8; Length 511;
Best Local Similarity 35.0%; Pred. No. 2, 3e-51;

Matches 168; Conservative 99; Mismatches 193; Indels 20; Gaps 7;

4 QNVVNGEMSSVNVQIETILSPIDDSLGFPV---AMTREEDVHMKAREAL---PAM 56
28 EFATVYNGEWTGTGAETVRSPIDSLIARVSLSMALSNRAVAAVYAGREIYDTG- 86
57 AALTYERAQYLHKAADIIERDEKEIATVLAKEISKAYNAVTEVVRTADLIRYAAEGI 116
87 -----EKLAEALFKVAELIRBSFDFTALVLDACKPISNARGVETATIERLEKTTMEFG 141
117 RLSTADEGGRKDASTGKLAIVIRQPVGIVLAIPYNPVNLSSGIAPALIGVNVWF 176
142 RLIGDIYIGDWSAEISGE-GIVKREPGVLAISPYNPPLFISTAKIVPALLAGNAVL 200

QY 177 KPPTGSGVGLVAKAFKAEAGLPAGVFNITGKSGEIGDYVHEBEVNFINFGSPVGO 236
DB 201 KPSGDPPLPPLLSRVLQNLGIPESYHILTVPGA-LMDSILDRIRAVFTGSTEVEB 259
QY 237 RIQKLAGMPIMLELGGKQAGIADADLDNAKQIVAGAYDYGSGORTAKRVLVEEV 296
DB 260 HILSMGIGIKFYHMEIGKQDPAVVLDDAPLEETVEKLVKGVNVSQGRCAIRILIEBGI 319
QY 297 ABELAKISENNAKLSVGP--DNATVTPVIDNSADPISLVDAROKAKELNEFRK 354
DB 320 YRDLKRELVAALSKIPENPLEDEDIMGPIINERSAEKIEEYVRDALKEKGAIVPLTGFR 379
QY 355 DGRLLTPGFD--HYTLPMKLAWEPPFGPIIIRVKAEEAVATANKSDGESSVPT 411
DB 380 KGAYVAPVLLBSREVLPGIRAFQEDVFGPLTLIVVSNEDAEVLANSSRFLDAVPS 439
QY 412 RFOKAFDIANKLEVGTVHINNKTGRGPNFPPLGKSGAGVQGIYRSIEAMTNKSV 471
DB 440 GDDSRARVAVARLEVGAVFINERPRHIGIYVPRGKMGDGIKSGIETITTTTKTIV 499

RESULT 52

ABU19008
ID ABU19008 standard; protein; 483 AA.

XX ABU19008;

XX 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #4535.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Bacillus anthracis.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342823P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haasebeck R, Ohlsen KU, Zyskind W,

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA22878.

XX Claim 25; SEQ ID NO 46932; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 483 AA;

Query Match 30.2%; Score 721.5; DB 6; Length 483;

Best Local Similarity 35.7%; Pred. No. 2,3e-51;

Matches 171; Conservative 82; Mismatches 203; Indels 23; Gaps 8;

QY 8 YVNGEMKSSVNOIEILSPIDSSLGFPVAMTREVVDHMKAGREALPAAALTYVERAY 67
DB 16 YINGEMITLQOEIVENNPAITKEIFATVPKGVTAKQAVDAHEAFKMSKLTADRAAK 75
QY 68 LHKADIIERDKEEIAITVLAKEISKAYNASVTEVVRTADLIYAAEBGRISTADEGK 127
DB 76 LKKWFTLLIDENKEEIEIAITKEQGRFAEALGEVYVANSFVWYABEGKRV-----YGHM 130
QY 128 MDASTGHKLAIVRRQVGVILAIAPVNVVNSGSKIAVPALGAVVWVFKPTGSGVSG 187
DB 131 IPASHPNKRIILVMKQPVGVMAAITPWNFPAMITKVPALAACTAAVVKPSQTPITL 190
QY 188 VLAKAFKAEAGLPAGVFNITGKSGEIGDYVHEBEVNFINFGSPVGOIRIKLAG--NR 245
DB 191 KLABLAHEADIPKGVINIVTGSKAKAIADTWMDGVRKVSFTGSEIGKELMAASAQTWK 250
QY 246 PIMLELGGKQAGIADADLDNAKQIVAGAYDYGSGORTAKRVLVVEVADDELAEKIS 305
DB 251 KVSLELGGHAPETVWMDADLDRAVEAVIGSKPRNAGQCTICTNRVFEVEVEAVEKEPE 310
QY 306 ENVAKLISVGDPF--DNATVTPVIDNSADPISLVDAROKAKELNEFRK----DGRLLT 360
DB 311 KAVGOLKRGDGRGDTTGGPIIDENAVSKVQDHIEDALQKGTULYGGQKVAELDGHFMQ 370
QY 361 P---GLFPHVTLDMKLAWEPPFGPIIIRVKAEEAVAIANKSDFGLOSVFTRDPK 417
DB 371 PTVIGLANDTMLCMN---EETFGVAPVAKFTVEBEVERANHTPYGLAAVIFTDISGA 427
QY 418 FDIANKLEVGTVHINNKTGRGPD--NPPPLGKSGAGVQGIYRSIEAMTNKSVLDM 474
DB 428 FOISALBEYGIILGNDGL---PSVAQAPFGFGKESGIGREGHFGIEBYLEIKYISLGL 483

RESULT 53

AAAR55135
ID AAR55135 standard; protein; 488 AA.

XX AAR55135;

XX 16-OCT-2003 (revised)

XX 18-JAN-1995 (first entry)

XX Heat resistant aldehyde dehydrogenase.

XX aldehyde dehydrogenase; heat resistant; alcohol production;

XX Bacillus stearothermophilus.

```

OS Geobacillus stearothermophilus.
XX Key Location/Qualifiers
FH CDS 1266..2732
FT /*tag= a
FT /product= "heat resistant aldehyde dehydrogenase"
XX
XX JP06113839-A.
XX
XX 26-APR-1994.
XX
XX 30-MAR-1992; 92JP-00074631.
XX
XX 30-MAR-1992; 92JP-00074631.
XX
XX (MARU-) MARUKAN SU KK.
XX
XX MPI: 1994-172737/21.
XX N-PSDB; AA065587.
XX
XX Heat-resistant aldehyde dehydrogenase from Bacillus stearothermophilus -
XX and corresp. DNA, useful for alcohol prodn.
XX
XX Claim 1; Page 8-10; 11pp; Japanese.
XX
XX This sequence shows an aldehyde dehydrogenase (ALDH) (see also AA065587),
XX that is heat resistant from Bacillus stearothermophilus and requires no
XX acetyl CoA. The fused protein of ALDH with an alcohol dehydrogenase
XX derived from a transformant which contains both structural genes can be
XX used in production of alcohol. (Updated on 16-OCT-2003 to standardise OS
XX field)
XX
XX Sequence 488 AA;
SQ
Query Match 30.2%; Score 721.5; DB 2; Length 488;
Best Local Similarity 34.4%; Pred. No. 2.4e-51;
Matches 166; Conservative 104; Mismatches 192; Indels 21; Gaps 9;
QY 3 KEYQYVNGEMKSV-NOIE-ILSPIDDSL-GFVPAMTREEDHAMKAREALPAMAL 59
DB 8 KTYFYVINGNMWTSVSNVPEPSINPANRDIYGVGRSLDEVDNEAVTANEAQISWMR 67
QY 60 TTY#QAQYIAKADIIERDKETATVLAKEISKANASVEVVRADLIRYAABSGIRUS 119
DB 68 SCVEGEGEYLKAAHLEQCLDIAETMTREMGKTLAEAKAETMRGVHILRYAAGEGAR-- 125
QY 120 TSADEGKMDASTGKHLAVIRROPVGIYLAIPYVPVNLSSGKIAPALIGNVVWFKRP 179
DB 126 ---KIGDVIPESSDSGLFTTTPVPLGAVVISPMNFPVALPIWKAPALVIGTVVLPKA 182
QY 180 TQGSVSGVLAKAFAPAGVFTTITGRSGEIGDYIEHSEVNFINFSTPVGQIRIG 239
DB 183 SFTAVTAKAVIECFHEAGFPKGVVNMVVGSGVGGGINHDPIDVFTTGSNTYKQVG 242
QY 240 KLAAGR--PIMELGSKAGIYLADADLNNAKQIYAGVYDSGQCTAIKRYLVVEEVA 297
DB 243 RAAPFRGAKYQLEMGGKKNPVIKADDLIAVEGTISGIRSTGQCTATSTVFIEREYV 302
QY 298 DLAEKISENVAKLSVGDPFDMAT-VTPVIDNSADFIESIYVVDARQKAKELNEFKR-- 354
DB 303 EFPKAKLIERVKQQLKINGLDAETWMPGCPASESQFTTIVSYIEKGSBEAKLIYGGNRCL 362
QY 355 -----DGRLLTPGLFDHVTLLDMKLAMEEPPFGPLPIIRYKDAEBAVAIANKSDPGLQSSV 409
DB 363 BEGLANGFVEPTEIFEDVDLQMTIAREELIFGPVLAIIQVDSIEBAIKLANDREYGLSASI 422
QY 410 FTRDQKAFDIANKLEVGTVHINNTKGRPDNFPPIGLK--GSGAGVQIRYSIEAMTV 467
DB 423 YKKNIGMLAEFTKIDIEAGLIKVNAETAGVEFOAPFGMGKSSSHSREOG-QAIEFTTSI 481
QY 468 KSI 470
DB 482 KTV 484

```

```

RESULT 54
ABB47807
ID ABB47807 standard; protein; 488 AA.
XX
XX ABB47807;
XX
XX 05-FEB-2002 (first entry)
XX
XX Listeria monocytogenes protein #511.
XX
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX vitamin B12; bacterial infection; disease.
XX
XX Listeria monocytogenes.
XX
XX WO200177335-A2.
XX
XX 18-OCT-2001.
XX
XX 11-APR-2001; 2001WO-FR001118.
XX
XX 11-APR-2000; 2000FR-00004629.
XX
XX (INSP) INST PASTEUR.
XX
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fajni H, Dehoux P,
XX Dussurgeat O, Chetoui F, Nedjari H, Glaser P, Kunst F, Coessart P,
XX Daniele U, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA,
XX Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,
XX Chakraborty T, Domann B, Hain T, Berche P, Charbit A, Durant L,
XX Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
XX Madueno E, De Pablo B, Wehlund J, Kaerst U, Entian K, Hauf J,
XX Rose M, Voss H;
XX
XX MPI: 2002-010914/01.
XX
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and related
XX polypeptides.
XX
XX Claim 6; SEQ ID NO 512; 192pp; French.
XX
XX The present invention relates to the genome sequence of Listeria
XX monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX it are useful for selecting probes and primers for detecting genes in L.
XX monocytogenes and related organisms, and for studying genetic
XX polymorphisms and other genomes. The present sequence is a protein
XX encoded by the genome sequence of the present invention. Proteins
XX expressed from the genome sequence are useful for raising specific
XX antibodies, identification of L. monocytogenes and related organisms, and
XX for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX B12. The genome sequence and proteins encoded by it are also useful for
XX selecting compounds that regulate gene expression and cell replication
XX and modulate L. monocytogenes-related diseases. In addition, the genome
XX sequence and proteins encoded by it are useful in pharmaceutical and
XX vaccines compositions for the treatment or prevention of infections by L.
XX monocytogenes and related organisms. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 488 AA;
SQ
Query Match 30.0%; Score 718; DB 5; Length 488;
Best Local Similarity 33.8%; Pred. No. 4.6e-51;
Matches 161; Conservative 96; Mismatches 201; Indels 18; Gaps 6;
QY 8 YVNGEMKSVNO--TEIUSPIDDSLGFPVPMTREEDHAMKAGREALPAMALTVYRA 65
DB 16 FINGKMTGDNKERTQIVNPANGVYIAKIAQAGSEYTKAIKAKADAPPDMAKMLADRV 75

```


[illegible]

RESULT 55
 ID ABU32914 standard; protein, 488 AA.
 XX
 AC ABU32914;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #18441.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS *Listeria monocytogenes*.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342823P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA36784.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 60838; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid

encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
 the target prokaryotic essential genes. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

Query Match	30.0%	Score 718;	DB 6;	Length 488;
Best Local Similarity	33.8%	Pred. No. 4.6e-51;		
Matches 161;	Conservative 96;	Mismatches 201;	Indels 18;	Gaps 6
QY	8	YVNGEWKSVNQ--IEILSPIDSSIGVPMPTREEVHAKAGREALPAMALTYVERA	65	
DB	16	FINGKMTDGNKERTKIDVNPANGVIATIAQAGSETKAIKLAADAPDMAKMELADRV	75	
QY	66	QYLKKAADIIRDKEEIAITVLAKETISKAYNASVTEVVRTADLIRYAAEBGIRLSTSADEG	125	
DB	76	KILKHADIAMEEKADTLAKIMTLEQKPLEKESKGEVLTVGVENFRFAAEEARL-----YG	130	
QY	126	GKMDASTGKLAVIRBOPGVIVLAAIPNPNVNLGSKIAPALIGANNVMPFPYQGSVS	185	
DB	131	ETIPPNNHAF-IVKKOPIGVVAALTFFNPFQGWTRKDLAPALATGNTIVLPSGDTPLS	189	
QY	186	GLVLAKAFAEAGLPAGVFNITTGSGSEIGDYIVEHEEVNFINFTGSTVVGORIGLAG--	243	
DB	190	ALAIPEIFEELAGLTKGVANIVMGSSKEIGELTLTSDVRKILFTGSTVVGQTLFPGQSDT	249	
QY	244	MRPIMLEIGSGDAGIVLADADLDNNAKOIVAGAVIDYSGQRCTAIRVLVBEVADBLAK	303	
DB	250	LKKSILEIGGAPFIVFPDDANIMDAVNDLVAKFPANNOQVCSPPRIIVAKEIKKEFTYA	309	
QY	304	ISENVAKISVGDPPDNNAVTPVVIDNSADPIESLIVVDAROKAGKELNEFKR-----DG	356	
DB	310	LVAKVQKLVNGSGDDVNVGPLIREDAIDDKDIOKATTEGAKVLTGGGRLTGSDYDKG	369	
QY	357	RLTLTPGLFDHTYLLDMKLAMEBPFQILPIIRVKDAEBAVAIAANKSDFGIQQSVFTTRFOX	416	
DB	370	NFYKPTVLDVNTKRMDDIYYEELFGVIVPLIFETEDALIMANDSEFGIASFYTKDLAR	429	
QY	417	AFDIANKLEVGTVHINNTKGGPDPNFPFLGKSGGAGYQGIIRYISEAMTNVYSITVL	472	
DB	430	VEKGAALAEYGVAGNEIAISNPF-PRGVGKHSQFGRGNHYGMEBYIQVKEFINTL	484	

RESULT 56

ADS29166

ADS29166 standard; protein, 493 AA.

AC

ADS29166;

XX

PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 18593; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 480 AA;
 XX
 Query Match 29.8%; Score 713.5; DB 8; Length 480;
 Best Local Similarity 35.0%; Pred. No. 1.1e-50;
 Matches 165; Conservative 86; Mismatches 210; Indels 11; Gaps 6;
 XX
 QY 8 YVNGEKKSSVNOJELISPIDSSLGFPVPMTRREVDHAKMGREALPAMALVVYRAQY 67
 DB 13 FVAGEVWSDEPVPVSDLDAGGTFASVAADTSDAEALSAATGVAHDLRETTVPVREV 72
 QY 68 LHKAAIIERKEEITVLAKEISKAVNSVTEVTRADLRVAAEGIRLSTADGCK 127
 DB 73 LESIADGIRREBELAEVIREAGKPISSARGVQSAAEFPDAVGE-LRLTGEVRTG- 130
 QY 128 MDASTGHR--LAVIRROPVGIATAPVNTVNLSSGKIAPALIGVNVMEKPTQSSVS 185
 DB 131 --TTAGHEBQVAIVKEHEPMGTVCITPVNPLSTMALQVAPALAGNAVIVKPSKTPIS 188
 QY 186 GVLAKAFANAGIPAGVFNITRGERSEIGDIYHEHEVNTINTGSPVQORIGKLAGRK 245
 DB 189 GALLADIAADAGLPDGAIVNFPGEBSVITGDLPLASDARVDIAIMTSSGAGEHVAROSGIT 248
 QY 246 PIVLEIGSKDAGIVLADADLDNNAKQIVAGAYDSGQRCFAIKRVIVVEEVADELAKIS 305
 DB 249 RLIMEIGGNAPALVFEDDADDAADAATAGSLKYAGRCGAVRVLAAHEVHDELVRID 308
 QY 306 ENYAKLSVGDPED-NATVTEVIDNSADFTIESLVVDAROKAGELEFKR--DG-RLLT 360
 DB 309 DANAEMSIGDLFTDTTLGPLVVSADQDWVAELVDAVDRGATVVGGERHVEDGVHYVE 368
 QY 361 PGFPHVTLDMKLAMEEPFPIPIIRYKDAEAAVAILANKSDGLQSSVTRPQAFDI 420
 DB 369 PTLIAVPRDARVDEQFGVPCAVTVTDEDAVTRANSEIALALAAVTTADHDAMRV 428
 QY 421 ANKLEVTGVHINKTKGSPNPPLGLKSGAGVQGIKRYVIEAMTVKSVTL 472
 DB 429 AERVNGAVRINQAPSHGLGVFPFGGNDASGIGREGLDSTIHFVNEKSIITL 480

RESULT 58
 ABUJ1620
 ID ABUJ1620 standard; protein; 482 AA.
 XX
 XX ABUJ1620;

XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #17147.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Klebsiella pneumoniae.
 XX
 PN WO20027183-A2.
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELITR-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyekind JW,
 PI Mail D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA35490.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX
 Claim 25; SEQ ID NO 59544; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WPI at
 CC ftp.wipo.int/pub/published_sequences
 XX
 SQ Sequence 482 AA;
 XX
 Query Match 29.7%; Score 711; DB 6; Length 482;
 Best Local Similarity 34.1%; Pred. No. 1.7e-50;
 Matches 163; Conservative 90; Mismatches 209; Indels 16; Gaps 7;

PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 PI WPI; 2004-061375/06.
 XX
 XX
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PS Claim 1, SEQ ID NO 18711, 122pp; English.
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX
 SQ Sequence 489 AA;
 Query Match 29.7%; Score 710.5; DB 8; Length 489;
 Best Local Similarity 35.9%; Pred. No. 2e-50;
 Matches 167; Conservative 84; Mismatches 199; Indels 15; Gaps 6;
 QY 8 YNNGEM--KSSVNOQIILSPIDSSIGFVPMTRFEEVDHAKRGREALPAMALTYVERA 65
 DB 15 YIDGEVPPGDRGTLEVDPPSTELFTEIPAGTTDDVAAYEAADAQEMADTPPOERA 74
 QY 66 QYIHKRAADIIEERDKETIATVYLAKEISKANASTVEVRRADLRVAEGIRLSTADSG 125
 DB 75 RIVAAASEILASHRDVLDLATESGTTTKGAEEFASVGLIDEAASPPTRKN-----G 129
 QY 126 GKMDASTGKHLAVIRQPGVIVLAIPVYVNVNLSGSKTAPALIGNVMPFPTGSGVS 185
 DB 130 DHKQSVBEKENIVAREPQGVGIIISPMFPLNLSRAVAPAAAGNVYLRPASTPTPT 189
 QY 186 -GVYAKAPAEAGLPAQVFNITIGRSGEIGDYVEHEEVNFINTSGTSPVQRIGTLAQ 244
 DB 190 GGLTIKVEAAGLPGGVFNVTGKSGEIGDRLAGPHADVISFTGSTAIGRVAAATAGB 249
 QY 245 RPIIM--LEIGKADAGIVLADADLDNAAKQIVAGATVYSGRCATKRVLVVEVADLAE 302
 DB 250 NLALPMEIGGNVHVLTEBDADDDVDGVSFSLHSQICISIRHVVHBEVYDEYVE 309
 QY 303 KISENVAKISVGPDPFNATVT--PVIDNSADFIESTLVVARQKAKELNFKRDLRTP 361
 DB 310 RLTERAALPTGTAHADATVYVGIIDESQRDQILDVITERTVADGALTEGSHDGLVYEP 369
 QY 362 GLFDHVTLLDMKAMEEPFGPILPIIVKQAEAVAIANKSDFGLQSVFTTRDFOKADIA 421
 DB 370 TVLSDATNTNMAAACHNHFPGVAVPIPFSDDEDAVEIANSTEVGLSGSVAGDTGRARDIA 429
 QY 422 NKLVEGTVAHNNKTGPGDNFPLGLKSGAGVQGIIRISLEMTN 466

DB 430 DRIDTGMTHINDQPNQGEHLPFGGYKNSGLG-----RVNGEYLVN 470
 RESULT 61
 ID ADS24538
 AC ADS24538 standard; protein; 473 AA.
 AC ADS24538;
 DT 02-DEC-2004 (first entry)
 DT XX
 DE Bacterial polypeptide #13571.
 DE XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 KW Bacteria.
 OS US2003233675-A1.
 PN 18-DEC-2003.
 PD 20-FEB-2003; 2003US-00369493.
 PF 21-FEB-2002; 2002US-0360039P.
 PR
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 PI WPI; 2004-061375/06.
 DR New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PT
 PS Claim 1, SEQ ID NO 13571, 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 CC

Sequence 473 AA:

Query Match	29.5%; Score 705.5; DB 8; Length 473;
-------------	---------------------------------------

Matches 167; Conservative 94; Mismatches 194; Indels 27

8 YVNGEWKSSVN--QIEILSPIDDSLGFPAMTREEDVDHAKAGREALPAWALTYERA 65

[illegible]

DR N-PSDB; ABA97235

PT New nucleic acid fragments encoding bacterial toluene monooxygenase
PT enzyme pathway, useful for isolating genes encoding proteins from the
PT same or other microbial species, and for producing para-hydroxybenzoate

Claim 4; Page 67-68; 97pp; English

CC The invention relates to an isolated nucleic acid fragment encoding a
CC bacterial toluene monooxygenase enzyme pathway. The nucleic acid
CC fragments of the invention may be used to isolate genes encoding proteins
CC from the same or other microbial species. Bacterial strains transformed
CC with the *pCreSol* utilizing (*pcu*) genes are useful for producing para-
CC hydroxybenzoate (PHBA) which can be used for synthesizing liquid crystal
CC polymers (LCP). The current sequence represents the *PcuC* amino acid
CC sequence, which has the enzyme activity of PHBAO. (Updated on 29-AUG-2003)
CC to standardise OS field)

Query Match 29.5%; Score 705.5; DB 5; Length 491;

Matches 167; Conservative 94; Mismatches 198; Indels 25; Gaps 8;

5 YONY-----VNGEMK--SSVNQIEILSPIDSSLGFPAMTREEDVDHAMKAGREALPAWAA 58

```

Db      8 YONLEIQPLAQOMRAGSSGRPLEVPDPYNDLELLRILASREDUDAAVRRKARDOSREMAT 67
Oy      59 LUYEBAOYLHNKRAADIERDKKEIATVLAKESIKAYNAVSVEVRTADLLRYAAEEGIRL 118
Db      68 TAAEEBAARVLLEAVKIFEDERREIEIDWIRSSGS-----TRIKQIEMGAAARITL 118
Oy      119 STSDEG---GKMDAST-GHKLAVIROPOGIVLALAPYNYPNVLSGSKTAPALIGENVV 174
Db      119 ESASLPNRYVHGRIIASNISGKESRYYRAPLGVIGISFPWNPPLHILTRARSLAPALACNAV 178
Oy      175 MFAPRPOQSASV-GLVILAKAPAGAPACVFNVTITGSGEIEDDYVEHEBVNFIFNTGTP 233
Db      179 VVKEPADDTPIITGGILLARIEERAGIPAGVLSVYVSGAEICDAVHEHPVALLISFTSTQ 238
Oy      234 VGORICKLAG---KRPIMLEIGSKDAGIVLADADLDLNAKQIVAGAYDYGORCTAIKR 289
Db      239 VGNRIIRISAGSEHLKTHVALIELGNSPVLVLADADVEOAVNAAVGKFLHQGQICVAINR 298
Oy      290 VLVEVEAEBALEKTSENVAKLSVCDPFPDNAVTV--PVIDDMSADVFISLVVDAROKAKE 348
Db      299 IIVGQPLEDFTRRPFVEYRKALPYDDBPSKPGTVGVPVINAQGLAKKATATAAEABATL 358
Oy      349 LNEFFKDGRLITPGLFDHVTLDMKLAAEPPGILPIIRVKDAEBAVALANKSDPGLQSS 408
Db      359 ILGGEPOGVMPHPHVGNTADMEIAREEIEGPIVIGIOSADADHALELANSSEYGLSSA 418
Oy      409 VFPKRDOKAFDLANKLEVGTVIINKKTORGPDNPPFLGLKSGSAGVGOGIRISIAMTNVK 468
Db      419 VFTASLERGVQFARRKHAGMTHVNDIYPNDEBNAPFGGEKNSGLGRNGDMALIEFFTDDH 478
Oy      469 SIYV 472
Db      479 WITL 482

```

```

XX OS Pseudomonas mendocina.
XX PN US2002151003-A1.
XX PD 17-OCT-2002.
XX PF 28-NOV-2001; 2001US-00997664.
XX PR 01-JUN-2000; 2000US-00585174.
XX PA (BENB/) BEN-BASSAT A.
XX PA (CATT/) CATERMOLE M.
XX PA (GATE/) GATENBY A. A.
XX PA (GIBS/) GIBSON K. J.
XX PA (RAMO/) RAMOS-GONZALEZ M. I.
XX PA (SARI/) SARIASLANI S.
XX PI Ben-Bassat A, Catermole M, Gatenby AA, Gibson KJ;
XX PI Ramos-Gonzalez M, Ramos JI, Sariaslani S;
XX DR MPI; 2003-428753/40.
XX DR N-PSDB; ACA92456.
XX PT Novel nucleic acid sequence encoding Tmost polypeptides, useful for
XX PT obtaining a nucleic acid fragment encoding Tmost polypeptides.
XX PS Example 1; Page 27-29; 65pp; English.
XX CC The invention describes an isolated nucleic acid fragment (I) of
XX CC Pseudomonas mendocina KR-1, comprising a 4821 nucleotide sequence and
XX CC encoding a Tmost polypeptide (referring to toluene monooxygenase
XX CC polypeptides, Tmos comprising 973 amino acids (a.a.s) having at least 83%
XX CC identity and Tmost comprising 220 a.a.s having at least 85% identity based
XX CC on the Smith-Waterman method of alignment with a sequence). (I) is useful
XX CC for obtaining a nucleic acid fragment encoding a Tmost polypeptide, by
XX CC probing a genomic library with (I), selecting for a DNA clone that
XX CC hybridizes with (I), and sequencing the genomic fragment that comprises
XX CC the identified clone, where the sequenced genomic fragment encodes Tmost
XX CC polypeptide. The polypeptide encoded by (I) is useful as a tool to
XX CC mediate expression of the catabolic two genes and para-hydroxybenzoate
XX CC (PHBA) production in any organism that does not possess (I). This is the
XX CC amino acid sequence of a protein encoded by the Pseudomonas mendocina p-
XX CC cresol utilizing (pcu) gene operon the proteins encoded by which are
XX CC involved in the toluene degradation pathway (TMO)
XX SQ Sequence 491 AA;
Query Match 29.5%; Score 705.5; DB 6; Length 491;
Best Local Similarity 34.5%; Pred. No. 5.2e-50;
Matches 167; Conservative 94; Mismatches 198; Indels 25; Gaps 8;
QY 5 YONY-----NGEKK--SSVNOQIEILSPIDSSLGFPNMTREVDNMRGRALPAMMA 58
DB 8 YONLEIQPIAGQGRAGSSGRPLEVDPYDELRLRLALSRDLDAYRKARDSQREWAT 67
QY 59 LTVYERAOYLHKADIIERDKERIATVLAKEISKAVNASVTEVFRADILRYAEEGIRL 118
DB 68 TAPAEARVULEVVKIFDRRERIIMTIRESS-----TRKAQIEMWAAARATL 118
QY 119 STSADG---GKMDAST-GHKLAVIRPQVIGLAIAPYNYVNLGSKTIAPALIGNAV 174
DB 119 ESSSLERVYGRILIANISGKESRVRAPGVISPMFPIHLTRASIALPALALGNV 178
QY 175 MFAPPTQGSVS-GLVYAKAPAEAGLPAQVNTITGSGSEIGDVIYHEEVENFNTGSP 233
DB 179 VVKPASDPTPTGGLLARIFEELGRLPAGVLSVVGSAELGDFVHPVPALISFTGSIQ 238
QY 234 VGGRIKLAG---MEPIMLELGGKAGIVLADADDLNNAKQIVAGAYDYSGQRCTAIR 289
DB 239 VGNIGRIASGGEHLKGVALLLEGGNSPFPVLLADADVGAVNAVVKFLHQGGICMAINR 298

```

```

QY 290 VLVEEVADELAEKISENVAKLSVGPPDNATVTT-PVIDNSAGDFIESLWVDAROKAKE 348
DB 299 LVEQPLEDFTRRVERVKALPYGDPKRGTVGEPVINAQOLAKUEKIAIRAKEGATL 358
QY 349 LNEFKRGRILTPGLFDHVTLMKLAMEEPFGILPIIRKDAEEAVALANKSDPGLQSS 408
DB 359 LIGGEQGVNMPPHFGVNTADMEIAREEIPPLVIGIQSARAHEHLAEANSSEYGLSCA 418
QY 409 VFTDRDQKAFDIANKLEVGTVHINKTGRGPDNPFPPLGLKSGAGVQGIKYSIEAMTNV 468
DB 419 VFTASLERGVQFARRIHAGMTHVNDIPVDEBNAPFGGERKNSGLGRFNGDMAIEEFTTDH 478
QY 469 SIYL 472
DB 479 WITL 482
RESULT 64
ABU40285
ID ABU40285 standard; protein; 480 AA.
XX AC ABU40285;
XX AC ABU40285;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #25812.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Pseudomonas putida.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyckind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR MPI; 2003-029926/02.
XX DR N-PSDB; ACA44155.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 68209; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 623 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an

```

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 480 AA;
 SQ
 Query Match 29.4%; Score 704; DB 6; Length 480;
 Best Local Similarity 34.0%; Pred. No. 6,7e-50;
 Matches 162; Conservative 91; Mismatches 209; Indels 14; Gaps 7;
 QY 6 QYVNGEMKSSVNV--QIEILSPIDDSLGFPVPMTEEDVHMKAREALPMAALTVE 63
 DB 12 QYVNGEMWDADNGQTITVNPATGSEVIGTVPKMGTAETRRALAEADKALPWARLTAK 71
 QY 64 RAQYLHKADIIERDKEEATVLAKEISKAVNASTVEVTRADLIRYAAEBSGIRLSTSD 123
 DB 72 REAKLRWFELMIENDDLARLMTTEQGRPLAEKEIYAAASFIMWPEAEKRIYGDIT 131
 QY 124 EGGKMDASTGHHKLAIVIRQPGVIALADYNTPVNLSSKIALPALIGNVNFKPTQGS 183
 DB 132 PGHQP-----KRLIYIKOPIGVTAITPMNFPMAITRKAGPALAAGCTMVLKPAQT 186
 QY 184 VSGVLAKAFABAGLPAGVFNITTGSGSEIGDIYVHEBVNFINTGSPVQGRICK--L 241
 DB 187 YALALVELAHRGIPAGVLSVTSAGVGELTGNLSVRKLSFGSTGIEGQLMECA 246
 QY 242 AGWRPIMLEBGKADGIVADADLDNAAQIVAGADYSGORCTAIXRLVVEEVADELA 301
 DB 247 KOIKKVSLEGGNAPRTIVDDDLDAVAGAIISKRRNGQTCVCAKRIYVQDGVYDARA 306
 QY 302 EKISENVAKLSVGPDPDAITV--PVTDNSADFIESLVVDARQKAKEINFKR--DGRLL 359
 DB 307 EKLAATAVAKLIGNGIEEGTITGPLLDGKAVAKVQHIDAVASKAKVSGGKLEGNFF 366
 QY 360 TGLPDIHVLDMKLAWEBSGPIPLIRYKDAEENAVANKSPFGSQSVFTRDPQKAD 419
 DB 367 EPTILVDVEKTAIVAKAEBETFGPLAPLFRFDEAEVITAMSNDEFEGLASTYFVADMSRVR 426
 QY 420 IANKLEVGTVAHINNKTR--GPDNFPPLGKSGAGVQGIKRSIEANTVKSIVLDM 474
 DB 427 VAAALEYGVNGIN--TGLISNEVAPRGKIKASGLGREGSKYGEDVLEIKYLCISV 480

OS Bacteria.
 XX US2003233675-A1.
 PN 18-DEC-2003.
 PD 20-FEB-2003; 2003US-00369493.
 PF 21-FEB-2002; 2002US-0360039P.
 PR
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PS Claim 1; SEQ ID NO 17169; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 SQ Sequence 468 AA;
 Query Match 29.1%; Score 697.5; DB 8; Length 468;
 Best Local Similarity 33.4%; Pred. No. 2.3e-49;
 Matches 158; Conservative 93; Mismatches 209; Indels 13; Gaps 5;
 QY 8 YVNGEMKSSVNVQIEILSPIDDSLGFPVPMTEEDVHMKAREALPMAALTVEYRQY 67
 DB 4 YVNGEMKSSGKTLVDVNPATGSEVITVPKMGTAETRRALAEADKALPWARLTAK 63
 QY 68 LHKADIIERDKEEATVLAKEISKAVNASTVEVTRADLIRYAAEBSGIRLSTSD 127
 DB 64 LMRWFQILDEQODELGEIMTEQGRPLAEKEIYAAASFIMWPEAEKRIYGDIT 118
 QY 128 MDASTGHHKLAIVIRQPGVIALADYNTPVNLSSKIALPALIGNVNFKPTQGS 187
 DB 119 IPASAINKRIIVQKQPGVIAIATPMNFPMAITRKAPALAAAGCTAIVKPAEQTPRAL 178
 QY 168 VLAKAFABAGLPAGVFNITTGSGSEIGDIYVHEBVNFINTGSPVQGRICK--GLA 245
 DB 179 KLAQLAEBAGIPAGVLSVTSAGVGELTGNLSVRKLSFGSTGIEGQLMRGAQYTK 238

[illegible]

Query Match	29.0%	Score 694	DB 8	Length 473
Best Local Similarity	32.6%	Pred. No. 4,56-49		
Matches 157	Conservative 102	Mismatches 198	Indels 24	Gaps 8
XX	Sequence 473 AA			
CC	tolerance to herbicides, extreme osmotic conditions, pathogens or pests,			
CC	increased resistance to plant disease, better growth rate by modification			
CC	of the cell cycle pathway with plant growth regulators, increased rate of			
CC	homologous recombination, modified seed oil or protein yield and/or			
CC	content, improved yield by modification of carbohydrate, nitrogen or			
CC	phosphorus use and/or uptake, by modification of photosynthesis or by			
CC	providing improved plant growth and development under at least one stress			
CC	condition, improved lignin production or improved galactomannan			
CC	production. This sequence represents a bacterial polypeptide used in the			
CC	scope of the invention. Note: The sequence data for this patent did not			
CC	form part of the printed specification but was obtained in electronic			
CC	format from USPTO at seqdata.uspto.gov/sequence.html.			
XX				
YY	8 YVNGEWKSVNQIEILS--PIDSSL-GFVPAMTRREVDHAMKAGREALPAMAALTYER 64			
DB	1 YINGQWLNAAETETTLNSHNPNADKSEIVATFPRSGADVDAVAABAQYSGMRKVPAPAR 60			
YY	65 AQYIHKAADITERDKEEIAIVLAKESIKAYNAVTEVYKPADLIRYAEGCIRLSTADE 124			
DB	61 AEYIFRVOGELLQHKKEIAQLISREMGKPLTEARGDVQEGVDCAFYSAGGGRLLF----- 115			
YY	125 GGRNDASTGHKLANIYRQRPVGIVALIAPNYVPVNLSSGSKAPALIGGNVVMFPPTQGSV 164			
DB	116 GQTPPEMNPENFATMTMPPIGVCALLTPMNPVPAALPCWKAMPALVCGNTVILRPADTPA 175			
YY	185 SGLVLAKAFAPAGLPAQVFTTITRSGEIDYIEHEEVAFINFTGSTPVGORIGKLGM 244			
DB	176 CATLIEIFFAAGLPPEVINLVHGVSEGVKALVEHNIDLVFTSSAAGVAYGEICGR 235			
YY	245 --RPIMLELGKDAIGIVLADADLDNAAQIVAGAYDYSQGRCTAIKEVLVEEVADELAE 302			
DB	236 THKVCLEMGCKNAQVVMEDADLELDGALMGAFGTGRCRTATSRLLIHRDICEKETTT 295			
YY	303 KISENVAKLSVGDFFDNAT--VTPIIDNSADFIESLVVDNRQGAKEI--NEFQRDRL- 358			
DB	296 MLRERTQQLRIGAGTEPEPTDIGPIINRQQRVHEVYNNIAREBAKLLIGGEIATEGOLQ 355			
YY	359 ----LTGGLFPHVTLDMKLAMEEPPGFBILPIIRKDAEEVAALANKSDPGLGSSVFRDF 414			
DB	356 QGYFGQPIILDNTVPQKRVAREEIFGVVALLIVSTTEBAIALINDPKYGLSSVYTRDI 415			
YY	415 QKAFDIANKLEVGTVAHINNKTGRGPNPPLGLKGSAGVQIGIRYSIEMNTKYSIVLDM 474			
DB	416 NRAVVAARDIEAGITTYINGPTIGAEVHLPGGVYKQTNGHR-----EAGTTALDVTEW 469			
YY	475 K 475			
DB	470 K 470			
RESULT 67				
ADN26088	ADN26088 standard; protein; 477 AA.			
XX	ADN26088;			
AC				
XX	02-DEC-2004 (first entry)			
DT				
XX	Bacterial polypeptide #8741.			
XX				
KM	Recombinant DNA construct; transformed plant; improved plant property;			
KM	cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;			
KM	pathogen tolerance; pest tolerance; plant disease resistance;			
KM	cell cycle pathway modification; plant growth regulator;			
KM	homologous recombination; seed oil yield; protein yield; carbohydrate;			
KM	nitrogen; photosynthesis; lignin; galactomannan;			
KM	bacterial polypeptide.			

```

XX OS Bacteria.
XX PN US2003233675-A1.
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX DR WPI; 2004-061375/06.
XX PT New recombinant DNA construct comprising a promoter positioned to provide
XX PT for expression of a polynucleotide encoding a polypeptide from a
XX PT microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 8741; 122pp; English.
XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plants with
XX CC properties.
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX CC increased resistance to plant disease, better growth rate by modification
XX CC of the cell cycle pathway with plant growth regulators, increased rate of
XX CC homologous recombination, modified seed oil or protein yield and/or
XX CC content, improved yield by modification of carbohydrate, nitrogen or
XX CC phosphorus use and/or uptake, by modification of photosynthesis or by
XX CC providing improved plant growth and development under at least one stress
XX CC condition, improved lignin production or improved galactomannan
XX CC production. This sequence represents a bacterial polypeptide used in the
XX CC scope of the invention. Note: The sequence data for this patent did not
XX CC form part of the printed specification but was obtained in electronic
XX CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 477 AA;
XX
XX Query Match 29.0%; Score 694; DB 8; Length 477;
XX Best Local Similarity 35.5%; Pred. No. 4,5e-49;
XX Matches 173; Conservative 83; Mismatches 197; Indels 34; Gaps 10;
XX
XX QY 5 YGNYNNGEKSSVNOJEIILSPIDSSLGVPATREVP-HAMKAREALPA----- 55
XX DB 3 YRNYIDGECWSASG---RTVDNVN---PADTRIVSHHAASDARDAAMAAAAA 54
XX
XX QY 56 --WAALTVERAQYLKKAADIERDKEEATVLAKEISKRYANVAVEVETADLIYAAE 113
XX DB 55 DGMKPIPKRAKILNDAAHLEANDTATLTREGRALNARBEVRSQOTLAFYAV 114
XX QY 114 EGIRLSTADBEGRKMDASTGHLKAVIRROPVIGVLAIAPIYVNVNLSGSKIADALIGNV 173
XX DB 115 EGOTFSGESYPPNDPD-----MLVYSLREPLAGVTVISPMWPFVSIARKIAPALITGNT 169
XX QY 174 WNEKPTQGSVSLVAKAFAEAGLPAQVNTITGSGSEIGDYIYHEEVPNTNFGSGTP 233
XX DB 170 VVFKSSDAPLPSGYRLAEAFVRAGIRKGVNLFLTGSAAAEVGPITVRESREAVSFTGSGT 229

```

```

XX QY 234 VGORIGKLAKMRP-IMLEIGGKDAIGYVLADDLDNAAKQIVAGADYSGORCTAIKRVLY 292
XX DB 230 AGEQIHKSVPMTTRTQMLGGKNPLIWMEDADLBRADVLAIVKGGSLSGQACTGSRILV 289
XX QY 293 VEEVADLEAKTISEVNAKLSVGDDPD-NATYTPVIDNSAPRISBLVYDAOK-----GA 346
XX DB 290 MEDVQAFTKELIVAKYKALKIGSGMTAGMDLGPATRKQLSTVLSYIAIGQEAATLGGG 349
XX QY 347 KEINFEK-RDGRLLRPGFHDVYTLDMKLAWEPEPGPIITIRYKQAEERAVAIANKSDRGL 405
XX DB 350 EQLSDGDRAHGYVAPVFTDTQSMRIAREEIFEVLALIEVDSYADALINQANDTEGL 409
XX QY 406 QSSVPTRDFOKAFDIANKLEVGTVHINNKTRGPDNPFELGKSGAGV--QGIRYSIEA 463
XX DB 410 SAAIATRPVYVHMDPTRDIESGTIVKINNTTGNLVNAPFGGLKRSSTPFGSGRAGLEF 469
XX QY 464 MTNFKSI 470
XX DB 470 YTOIKTV 476
XX
XX RESULT 68
XX ABU41623
XX ID ABU41623 standard; protein; 480 AA.
XX AC ABU41623;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #27150.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Pseudomonas syringae.
XX PN WO20027183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002MO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA45493.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 69547; 176pp; English.
XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 623 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of

```

CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (6)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
XX Sequence 480 AA;
SQ
Query Match 29.0%; Score 693; DB 6; Length 480;
Best Local Similarity 33.8%; Pred. No. 5.5e-49;
Matches 161; Conservative 89; Mismatches 212; Indels 14; Gaps 7;
QY 6 QNYVNGEM--KSSVNOIELSPIDSSLGFPAMTBEVDHAKGREALPMAALTYVE 63
DB 12 QAYINGQMDADGGGSIKVNPNATNELGTVPKGAETRRALIBADKALPARALITAKE 71
QY 64 RAQYLHKADIIERDKSEIATVLAKEISKAYNASVTEVTRADLIRYAEGRISTAD 123
DB 72 RGNKLRWFELMTENQDGLMTLEQKFLAEKGEITYAASFIEWFABEAKRYVDVY 131
QY 124 EGGKMASTGHKLAIVRROPVGIYLAIPNYPVNLSSGTAIALTGANNVMPKPGQS 183
DB 132 PGHQPD---KRLIVAKOPIGVTAATTPNNFPAMTRAGPALAGCTWVKEPASQTP 186
QY 184 VSGQLVLAKEFAEAGLPAGVFNITTGSGSEIGDYIVHEEVNPFINTGSPVPGORIGK--L 241
DB 187 FSLALALAEIARIGIPAGVSVTGTSGSGLGSELTNGPIRYKLSFTGSTEIGQWAMECA 246
QY 242 AGNRPMLEIGSGDAGIVLADADLNNAKQIVAGAYDSGQRCATIKRVLVEEVADELA 301
DB 247 KDIKKTSLSELGNAPFIVFPDADLDKRAVEGAMTSKYNNQTCVCANRIYVQGVDAFA 306
QY 302 EKISENVAKISVQDPF--DNATVTPVVDNADPISLIVNARQKAKELNEFKR--DGRLL 359
DB 307 EKKVAVAGKLIKNGLEDGITTPBLIDEKAVAVKEHIDAVSKGATVLTGNSLRSQFP 366
QY 360 TPGLFHVTLDMKLAMEEPGPILPITRVDAEAVANAKSPDGLQSVFTDFOKAPD 419
DB 367 EPIIVNVSDAIVARETTPGLAPLFRPKDBAALITANDTERGLASYVAQNMGRVFR 426
QY 420 IANKLEVGTVHINKTGRGPDNF--PVLGKSGAGVQGIIRYSIEAMTNVKSIVLDM 474
DB 427 VAAALEGVWGIN--TGLISNELAPFGIGIKSGIGREGSKYGIEDYLEIKYLCISV 480
RESULT 69
AB042152
ID AB042152 standard; protein; 480 AA.
XX
XX AB042152;
XX
XX 19-JUN-2003 (first entry)
DE Protein encoded by prokaryotic essential gene #27679.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS Pseudomonas syringae.

XX
PN MO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR MPI; 2003-029926/02.
XX
PS N-PSDB; ACA46022.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 70076; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 480 AA;
Query Match 29.0%; Score 693; DB 6; Length 480;
Best Local Similarity 33.8%; Pred. No. 5.5e-49;
Matches 161; Conservative 89; Mismatches 212; Indels 14; Gaps 7;
QY 6 QNYVNGEM--KSSVNOIELSPIDSSLGFPAMTBEVDHAKGREALPMAALTYVE 63
DB 12 QAYINGQMDADGGGSIKVNPNATNELGTVPKGAETRRALIBADKALPARALITAKE 71
QY 64 RAQYLHKADIIERDKSEIATVLAKEISKAYNASVTEVTRADLIRYAEGRISTAD 123
DB 72 RGNKLRWFELMTENQDGLMTLEQKFLAEKGEITYAASFIEWFABEAKRYVDVY 131

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway;
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 489 AA;

Query Match 28.8%; Score 690; DB 6; Length 489;

Best Local Similarity 33.4%; Pred. No. 1e-48;

Matches 161; Conservative 93; Mismatches 200; Indels 26; Gaps 8;

QY 8 YNNGWSSVNOEITSPIDSSLG---FVPMTRSEVDHAKAGREALPMAALTVE 63
 DB 19 YIDGKVVAGNPSI--PYDNSTGKTIISVPLGKREAEQADAEALPMSARTGKE 76
 QY 64 RAQVYHKRADIERDKEEIAATLAKESKAYNASVTEVRTDLRYAEEGRISTSD 123
 DB 77 RAAILKMKQIMMENQODLAVITMSEQKVPBAEGLIYVASFLEWFAEAKRI---- 131
 QY 124 EGGKMDASTGHTLAVIRQPGVILAIAPYNYVNLGSKIAIPALIGVNVWFKPTQGS 183
 DB 123 DDDIIQSPKGGGRWLWKQPIGVTAITPWNPPAAMITRKVPALAACTWVVKRQGP 191
 QY 184 VSGILVIAKAPADAGLPAGVFNITTGSGSIGDIYVEHEVNFINTGSTPVGQRIKLAG 243
 DB 192 LITALLAVIAGVAGVAGVQVVTGSSREIGALCESEVVRKLSFGSTEVGRTIMEQCA 251
 QY 244 --MRPIMELGGKMDAGIVLADADLDAKQIYAGADYGGQCTAKRVLVYVEVADELA 301
 DB 252 PTIKKLSLELGNAPRTVPDDADLDRAIDGILASKRYRNGQTCVCANRIYQAGYEEVA 311
 QY 302 EKISENVAKLSVGDPEP--NATVTPVIDNSADFIESTLVVDARQKAKEL--NEFKRDR 357
 DB 312 KRLVEKVNAMKVGDEGEQVTOGPLIDTAIVAKVOGHIDATRAHGAIVAGSEPHALGGS 371
 QY 368 LITGLFDHVTLDKLAWEPPGPILPIIRVDAEBAVALANKSDGLQSSVFTDRDQKA 417
 DB 372 FPGQFVVRDVTQSMFAVEETFGPVAPLFEKPSSEDEIVGMANNITFGLAAVFTRYARI 431
 QY 418 FPIANKLEVTGTHIN-----NKTGRGPDNFPPLGLKSGSAGQGRIRYSTEANTVKSITLV 472
 DB 432 KWVSEALEYIGVINTGIISNEVG-----PFGVQVQSGLGRBGSKYGIEDYIELKYLVCV 485
 QY 473 DM 474
 DB 486 DL 487

RESULT 73
 ADS22151
 ID ADS22151 standard; protein; 480 AA.

XX ADS22151;
 XX

DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #1184.

KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKER G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide

XX for expression of a polynucleotide encoding a polypeptide from a

XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 11184; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 480 AA;

Query Match 28.7%; Score 687; DB 8; Length 480;

Best Local Similarity 32.1%; Pred. No. 1.8e-48;

Matches 155; Conservative 105; Mismatches 207; Indels 16; Gaps 6;

QY 1 LTKQYQNVNGEWMK--SSVNOEITSPIDSSLGVPFPMTRSEVDHAKAGREALPMAA 58
 DB 3 MVQYKLFITGFBKDSSTGETFEDINPATYLENLATIQVAGABDVMAVEAABAGFRLWNE 62

XX 02-DEC-2004 (first entry)
DT Bacterial polypeptide #7037.
DE
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.
OS Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 7037; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 458 AA;

Query Match 28.5%; Score 682.5; DB 8; Length 458;
Best Local Similarity 36.1%; Pred. No. 3.9e-48;
Matches 168; Conservative 84; Mismatches 199; Indels 15; Gaps 7;

QY 11 GEMKSSVNOIEILSPIDSSSLGFVPMTRSEVDHAKAGREALPMALTYTERAQYLHK 70
DB 1 GRRVTRPRTLVDVDFYTGTVGPLAS---VDVRAAAFEYAAYQAOLTRYERSQILER 57

QY 71 AADIIERDKERIAVLAKEISKAVNASYTEVVRTADLIRYAABEGIRLSTADEGGRKDA 130
DB 58 AASIMBERTEEASDLISLESGLSKQDSRYEIGRVADVLFKFASTELR-----DDGQSPSC 112
QY 131 ST---GKTLAVI-RRQPV-GIVLAIAPIYNYVNLSSGKIAPALIGNVVMFKPPTQSSVS 185
DB 113 DLTPHGKRRRYFSQREPLAGVIAITPPHNMNOVAHKIAIAIATNNRVLKPKSEKPLS 172
QY 186 GLVLAKAFAEAGLPAGVNTLTGSGSEIGDYIVHEEBVNFNFPGSTPVGGRIGKLAMR 245
DB 173 ALVYLDVLYEAGLPAPMLQVLTGDPRETADLITPLAEVLFTGCVATIGKTIARAAYR 232
QY 246 PIMLEGGKADAGIYLADADLDNAAKQIVAGAYDYSQGRCTAIRKVLVVEEYADEIAEKIS 305
DB 233 RVLELGNDPLIYLDADLERRAATLAVQSGYSKNSGQRCITAVKMLVQKSVAADPTDLVV 292
QY 306 ENVAKLSVGDPD-NATVTPYIDNSADFTISLVVDARQKAKELNEKRGRLITPOLF 364
DB 293 EKTRAWTPGDPFDASSQWGTIVDVAALFPEARVNBAAVASGARLLTGNQRNGALYAPTVL 352
QY 365 DHVTLDMKLAWEBPPEPILPTIIRVKDAEVAIAINKSDPGLQSSVFTEDPQKAPDIANKL 424
DB 353 DGVDPSMTLYRBEETFGVPSPILITPTDLDARISNGTAFGLSSGVCTNRQALITRFINEL 412
QY 425 EVGTIVHINKTGRGPDNPFPLGLKSGAGV-QGIRYSIEAMTNVYS 469
DB 413 RVGTIVWEVPGYRIELTFPGIGKDSGLGYKEGVQAMKSFYNLKT 458

Search completed: August 28, 2005, 10:24:46
Job time : 177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2005, 10:26:35 ; Search time 162 Seconds
(without alignments)
1152.098 Million cell updates/sec

Title: US-09-868-195-12

Perfect score: 2393

Sequence: 1 LTRKEYQNYNGEMKSSVNDI.....GIRYSLEANTNKSTIVLDMK 475

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 80 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1812	75.7	474	16	US-10-474-776-667
2	1812	75.7	474	17	US-10-472-928-2228
3	1802	75.3	474	15	US-10-282-122A-73974
4	1700	71.0	475	15	US-10-282-122A-72316
5	1699	71.0	439	9	US-09-815-242-13316
6	1692	70.7	436	18	US-10-617-320-3804
7	1639	68.5	475	15	US-10-282-122A-74625
8	1397	58.4	486	15	US-10-282-122A-53543
9	1388.5	58.0	486	15	US-10-282-122A-52739
10	1383	57.8	482	15	US-10-282-122A-52073
11	1335	55.8	472	15	US-10-369-493-16536

12	1312	54.8	473	15	US-10-369-493-16664	Sequence 16664, A
13	1244	52.0	498	15	US-10-369-493-17301	Sequence 17301, A
14	1172.5	49.0	491	16	US-10-425-115-215964	Sequence 215964, A
15	1146.5	47.9	503	16	US-10-437-963-189014	Sequence 189014, A
16	1038.5	43.4	503	15	US-10-424-599-187423	Sequence 187423, A
17	1001.5	41.9	434	15	US-10-424-599-187419	Sequence 187419, A
18	868.5	36.3	475	15	US-10-282-122A-76821	Sequence 76821, A
19	786.5	32.9	492	14	US-10-268-518-4	Sequence 4, Appl1
20	786.5	32.9	492	16	US-10-391-364-28	Sequence 28, Appl1
21	777	32.5	493	9	US-09-823-901-9	Sequence 9, Appl1
22	777	32.5	493	14	US-10-175-666-21	Sequence 21, Appl1
23	777	32.5	493	16	US-10-776-871-21	Sequence 21, Appl1
24	775.5	32.4	492	14	US-10-172-585-13	Sequence 13, Appl1
25	763	31.9	488	15	US-10-369-493-23067	Sequence 23067, A
26	759.5	31.7	463	15	US-10-369-493-1073	Sequence 1073, A
27	741.5	31.0	455	15	US-10-369-493-1188	Sequence 1188, A
28	741.5	31.0	484	9	US-09-738-626-6430	Sequence 6430, Ap
29	741.5	31.0	484	16	US-10-781-014-228	Sequence 228, App
30	737	30.8	476	15	US-10-369-493-17	Sequence 17, Appl1
31	735	30.7	482	15	US-10-369-493-17249	Sequence 17249, A
32	728	30.4	470	15	US-10-369-493-16454	Sequence 16454, A
33	726	30.3	482	9	US-09-815-242-10264	Sequence 10264, A
34	726	30.3	482	15	US-10-369-493-23511	Sequence 23511, A
35	726	30.3	482	15	US-10-282-122A-56650	Sequence 56650, A
36	721.5	30.2	488	15	US-10-282-122A-60838	Sequence 60838, A
37	718	30.0	488	15	US-10-282-122A-60838	Sequence 60838, A
38	716.5	29.9	493	15	US-10-369-493-18199	Sequence 18199, A
39	713.5	29.8	480	15	US-10-369-493-18593	Sequence 18593, A
40	711	29.7	482	15	US-10-282-122A-59544	Sequence 59544, A
41	710.5	29.7	489	15	US-10-369-493-18711	Sequence 18711, A
42	705.5	29.5	473	15	US-10-369-493-13571	Sequence 13571, A
43	705.5	29.5	491	9	US-09-997-664-3	Sequence 3, Appl1
44	705.5	29.5	491	15	US-10-464-952-3	Sequence 3, Appl1
45	704	29.4	480	15	US-10-282-122A-68209	Sequence 68209, A
46	697.5	29.1	468	15	US-10-369-493-17169	Sequence 17169, A
47	694	29.0	473	15	US-10-369-493-18996	Sequence 18996, A
48	684	29.0	477	15	US-10-369-493-8741	Sequence 8741, Ap
49	683	29.0	480	15	US-10-282-122A-69547	Sequence 69547, A
50	683	29.0	480	15	US-10-282-122A-70076	Sequence 70076, A
51	682	28.9	480	15	US-10-282-122A-75217	Sequence 75217, A
52	681	28.9	478	15	US-10-369-493-13957	Sequence 13957, A
53	680	28.8	489	15	US-10-282-122A-51224	Sequence 51224, A
54	687	28.7	480	15	US-10-369-493-11184	Sequence 11184, A
55	686	28.7	467	15	US-10-369-493-8312	Sequence 8312, Ap
56	682.5	28.5	458	15	US-10-369-493-4282	Sequence 4282, Ap
57	682.5	28.5	485	15	US-10-369-493-7037	Sequence 7037, Ap
58	682.5	28.5	485	15	US-10-369-493-17279	Sequence 17279, A
59	682	28.5	482	9	US-09-815-242-14047	Sequence 14047, A
60	682	28.5	482	15	US-10-282-122A-75920	Sequence 75920, A
61	680.5	28.4	460	15	US-10-369-493-16681	Sequence 16681, A
62	677	28.3	537	15	US-10-424-599-162810	Sequence 162810, A
63	676.5	28.3	491	15	US-10-369-493-4453	Sequence 4453, Ap
64	675	28.2	481	15	US-10-369-493-12269	Sequence 12269, A
65	673	28.1	485	15	US-10-369-493-23361	Sequence 23361, A
66	672	28.1	474	14	US-10-273-051-12	Sequence 12, Appl1
67	672	28.1	474	15	US-10-129-518-12	Sequence 12, Appl1
68	671.5	28.1	499	14	US-10-156-761-11332	Sequence 11332, A
69	670.5	28.0	491	15	US-10-369-493-21333	Sequence 21333, A
70	669.5	28.0	470	15	US-10-369-493-16985	Sequence 16985, A
71	668	27.9	468	15	US-10-369-493-7213	Sequence 7213, Ap
72	668	27.9	483	9	US-09-815-242-5058	Sequence 5058, Ap
73	668	27.9	483	15	US-10-282-122A-43445	Sequence 43445, A
74	667.5	27.9	481	8	US-08-976-063C-28	Sequence 28, Appl1
75	667.5	27.9	481	10	US-09-750-996D-28	Sequence 28, Appl1
76	667	27.9	240	16	US-10-767-701-33400	Sequence 33400, A
77	666.5	27.9	524	15	US-10-369-493-23457	Sequence 23457, A
78	666	27.8	485	15	US-10-369-493-7145	Sequence 7145, Ap
79	665.5	27.8	490	15	US-10-369-493-13769	Sequence 13769, Ap
80	665	27.8	474	15	US-10-369-493-4390	Sequence 4390, Ap

ALIGNMENTS

```

; RESULT 2
; US-10-472-928-2228
; Sequence 2228, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926W0
; CURRENT APPLICATION NUMBER: US/10/0472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 2228

```

```

RESULT 3
US-10-282-122A-73974
; Sequence 73974, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Habelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: E01TA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
;

```

```

; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 73974
; LENGTH: 474
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73974

Query Match          75.3%; Score 1802; DB 15; Length 474;
Best Local Similarity 73.7%; Pred. No. 5 8e-136;
Matches 347; Conservative 58; Mismatches 66; Indels 0; Gaps 0;

QY 5 YQNVYNGEMKSSVNOIEILSPIDSSLGFPVPMPTREEDVHAMKAGREALPAMALTYVER 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 4 YQNVYNGEMKSSVNOIEILSPIDSSLGFPVPMPTREEDVHAMKAGREALPAMALTYVER 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 64 AAYLHAKTAALIERDKEEATVLAKEISKAYNASVTEVTRADLIRAAEBGRLSTASDE 124
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 125 GSKMDASTGHKLAVIRROPVGIYLAIAFYNYPNVLSGSKIAAPALIGNVVMFKPTQGSY 184
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 124 GGGFPAASAKKLAVVAREVGVILALAFNYPNVLSASKIAPALIGNVVMFKPTQGSY 183
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 185 SGVLAKAPAEAGLPAGVFNITIGRSGEIGDYIYEHEEVNFINTGSTPVGORIGKLAM 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 184 SGVLAKAPAEAGLPAGVFNITIGRSGEIGDYIYEHEEVNFINTGSTPVGORIGKLAM 243
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 245 RPIMELGKGDAGIVTADADLNNAAQIYAGAYDYGORCTAKRYLVVEEVADELAEKI 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 244 RPIMELGKGDALVIEDADLEHAAKOIYAGAFSYSGORCTAKRYLVVEEVADELAEKI 303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 305 SENVAKLSVGDPPDNATVTPVIDDNGADFIESLVVDAROKGAKELNEFRDGRILTPGF 364
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 304 QEBVSRLTIVGDPDNADITPVIDNASADFIWGLIEDAOKERQALTPIRBEGMLPVL 363
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 365 DHTVLDMLKLAEBEPGPILPIIRVKDAEVAIALANKSDFGLOSSVTRDPFOKAFDIANTL 424
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 364 DQVTKOMKXAMEBEPGPVLPPIIRVASVEEALIAFANSEFGLOSSVTRDPFOKAFELAEKL 423
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 425 EYGTVINNKTRGPGPNPFLIGKSGAGVQGIIRYSIEAMTNKSVILDMK 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 424 EYGTVINNKTRGPGPNPFLIGKSGAGVQGIIRYSIEAMTNKSVILDMK 474
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
US-10-282-122A-72316
; Sequence 72316; Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
```

```

; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 72316
; LENGTH: 475
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-10-282-122A-72316

Query Match          71.0%; Score 1700; DB 15; Length 475;
Best Local Similarity 66.9%; Pred. No. 9.1e-128;
Matches 318; Conservative 75; Mismatches 82; Indels 0; Gaps 0;

QY 1 LKREYQNVYNGEMKSSVNOIEILSPIDSSLGFPVPMPTREEDVHAMKAGREALPAMALTY 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MKQYKQNVYNGEMKSSVNOIEILSPIDSSLGFPVPMPTREEDVHAMKAGREALPAMALTY 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 YVERAQYLHKAADIIERDKEEATVLAKEISKAYNASVTEVTRADLIRYAAEBGRLST 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 YVERAQYLHKAADIIERDKEEATVLAKEISKAYNASVTEVTRADLIRYAAEBGRLST 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 SHADGGKMDASTGHKLAVIRROPVGIYLAIAFYNYPNVLSGSKIAAPALIGNVVMFKPT 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 SHADGGKMDASTGHKLAVIRROPVGIYLAIAFYNYPNVLSGSKIAAPALIGNVVMFKPT 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 OGSVSGVLAIAKAPAEAGLPAGVFNITIGRSGEIGDYIYEHEEVNFINTGSTPVGORIGK 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 OGSVSGVLAIAKAPAEAGLPAGVFNITIGRSGEIGDYIYEHEEVNFINTGSTPVGORIGK 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 LAGMPRIMLELGKDGAGIVTADADLNNAAQIYAGAYDYGORCTAKRYLVVEEVADEL 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 LAGMPRIMLELGKDGAGIVTADADLNNAAQIYAGAYDYGORCTAKRYLVVEEVADEL 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 AKEISENVAKLSVGDPPDNATVTPVIDDNGADFIESLVVDAROKGAKELNEFRDGRILTP 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 AKEISENVAKLSVGDPPDNATVTPVIDDNGADFIESLVVDAROKGAKELNEFRDGRILTP 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 PGLPHTVLDMLKLAEBEPGPILPIIRVKDAEVAIALANKSDFGLOSSVTRDPFOKAFDI 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 361 PGLPHTVLDMLKLAEBEPGPILPIIRVKDAEVAIALANKSDFGLOSSVTRDPFOKAFDI 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 ANKLEVGTVHINKTKRGPGPNPFLIGKSGAGVQGIIRYSIEAMTNKSVILDMK 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 421 ANKLEVGTVHINKTKRGPGPNPFLIGKSGAGVQGIIRYSIEAMTNKSVILDMK 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 5
US-09-815-242-13316
; Sequence 13316, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13316
; LENGTH: 439
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13316

Query Match 71.0%; Score 1699; DB 9; Length 439;
Best Local Similarity 74.5%; Pred. No. 9,7e-128;
Matches 327; Conservative 55; Mismatches 57; Indels 0; Gaps 0;

Qy 37 MTRREVDHAKGREALPMAALTYVERAOYLHKADIERDKBEIATVLAKISKAYNA 96
Db 1 MGTLEDEDMQARALPMPRALSAIERAAVYHKTAIERDKKIGTTLAKENAKGIRA 60
Qy 97 SVTETVTRADLIRYAAEBGIRLSTSDSGKMDASTGHTKLAVIRROPVGIIVLAIAPYNP 156
Db 61 AIGGVTRADLIRYAAEBGLRTGQMEGGFEAASKNKLAVRRPVGIVLAIAPNPVP 120
Qy 157 VNLSSGKIAPALIGNVVFKPPTQGSVGLYLAKFAPAAGLPAGVFNTITGRGSETIGY 216
Db 121 VNLSSASKIAPALIGNVVLMKPPTQGSISGLILAKAFEEAGIPAGVFNTITGRGSEIGDY 180
Qy 217 IYEHHEVNFINTGSPVGORIGKLAGMPPIMELEGCKAGIYLAADIDNNAKOIVAA 276
Db 181 IYEHHEVNFINTGSPVGERIGRLAGMPPIMELEGCKAPALVLEADIDNNAKOIVAA 240
Qy 277 YDSGGRCTAIRKRVILVEEVADELAEKISENVAKLSDVDPFDNAVTPVIIDNSADFISS 336
Db 241 FYSYGGRCTAIRKRVILVESVADKATLLQEVSKLVGSPFDNADITPVIIDNSADFIWG 300
Qy 337 LVVDAROKAKELINEKRDGRLLTTPGLPHVTLDMKLANEBEPGPILPITIRYKDAEAYA 396
Db 301 LLEDQOEKKAQALTPIKREGNLMVPLFPQVTKDMKCVAMBEPPGPVLPITIRVASYEAYA 360
Qy 397 IANKSDPGLQSSVTRPDPOKAPDIANKLFGVTYHINKKRGPDNFPPIGLGSGAGVGG 456
Db 361 FANSEBFGLOSSVTFNDPKKAFBIAEKLEVGIVHINKTKRGPDNFPPIGLGSGAGVGG 420

Qy 457 IRYSIEMANTVKSIVLDMK 475
Db 421 IKYSIEMANTVKSIVFDVK 439

RESULT 6
US-10-617-320-3804
; Sequence 3804, Application US/10617320
; Publication No. US20050136404A1
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stream and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/617,320
; FILING DATE: 10-Jul-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arcinieglo, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3804:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...436
; SEQUENCE DESCRIPTION: SEQ ID NO: 3804:
US-10-617-320-3804

Query Match 70.7%; Score 1692; DB 18; Length 436;
Best Local Similarity 74.9%; Pred. No. 3,5e-127;
Matches 326; Conservative 53; Mismatches 56; Indels 0; Gaps 0;

Qy 41 EYDHAHMKREALPMAALTYVERAOYLHKADIERDKBEIATVLAKISKAYNAVTE 100
Db 2 EADBSGQARALPMPRALSAIERAAVYHKTAIERDKKIGTTLAKENAKGIRAIGE 61
Qy 101 VRTADLIRYAAEBGIRLSTSDSGKMDASTGHTKLAVIRROPVGIIVLAIAPYNPVNL 160
Db 62 VRTADLIRYAAEBGLRTGQMEGGFEAASKNKLAVRRPVGIVLAIAPNPVNL 121
Qy 161 GSKIAPIIGNVVFKPPTQGSVGLYLAKFAPAAGLPAGVFNTITGRGSETIGYIEH 220

Db 122 ASKIAFALLAGNVVMEKPTQSSISGLLAKAFEEAGIPAGVFNITITGSGEIGDYIIEH 181
Qy 221 EEVNFINTGSIPTVGORIGLQAGMPIMEELGSKDAGIVLADDLNNAKQIVAGAYDYS 280
Db 182 KEVNFINTGSIPTIGRIGRLAGMPIMEELGSKDALVLEADLEHAAKQIVAGAFSYS 241
Qy 281 GORCTAIKRVLYVEEVADELAEKISNVAKLSVGDPPDNATVTPVDDNSADIEELVYD 340
Db 242 GORCTAIKRVLYVESVADKLATLLQSEVSCLTVGDPPDNADITPVIDNSADIEELVYD 301
Qy 341 AROKGAKEINEFGRDRLLTPGLFDHVTLLDMKLAEEEPGPIPIIRVKDAEEAVALANK 400
Db 302 AOKKEAOLTPIRREGNLLMPVFPDYTKDMKVAMEEPGPVPIIRVASEELIAFANE 361
Qy 401 SDGLOSSTFTDPOKAFDIANKLEVGVTHINNKTRGPDNPPLGLKSGAGVQIGIRYS 460
Db 362 SEFGLOSSVFTNPFKAFAIEAEKLEVGVTHINNKTRGPDNPPLGLKSGAGVQIGIRYS 421
Qy 461 IEAMTVKSIIVLDMK 475
Db 422 IEAMTVKSIIVFDVK 436

RESULT 7
US-10-282-122A-74625
Sequence 74625, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remainder of SEQ ID NOS: 78614
Remainder of SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 74625
LENGTH: 475
TYPE: PRT
ORGANISM: Streptococcus pyogenes
US-10-282-122A-74625

Query Match 68.5%; Score 1639; DB 15; Length 475;
Best Local Similarity 65.5%; Pred. No. 7, 2e-123;
Matches 311; Conservative 77; Mismatches 87; Indels 0; Gaps 0;
Qy 1 LTRKGVYONVNGEMKSVNOIEILSPIDSLGFPVPMTRREYDHAKAGREALPAWAALT 60
Db 1 MAKQYGNLVNGEMKLSSENITTYAPATGSELDSVPAMTQAEVDAYASAKKLSDBRALIS 60
Qy 61 VYERAPYHKAADIIERDEKEIATYLAKEISAYNNAVTEVVRTPADLLIYAAEEGIRLST 120
Db 61 VYERAPYHKAADIIYVRDAEKIATLSKEVAKGHKAAVSEVIRTAIIINYAAEEGIRMEG 120
Qy 121 SADEGSKMASYGHKLAIVRRQPVGIVLAAPYNPVNLSGSKIAFALLAGNVVMEKPT 180
Db 121 EYLEGSGFPAASKKKAIAVRRPVGIVLAISPNNPVNLASGKIAPALLAGNVVMEKPT 180
Qy 181 QGSVSGLVLAAPAEAGLPAAGVFNITITGSGEIGDYIIEEVEVNFINTGSIPTVGORIG 240
Db 181 QGSISGLLAEAPAEAGIPAGVFNITITGSGVIGDYIIEHVAEVSFINTGSIPTIGRIG 240
Qy 241 LAGMRPIMEELGSKDAGIVLADDLNNAKQIVAGAYDYSGORCTAIKRVLYVEEVADEL 300
Db 241 LAGMRPIMEELGSKDASAIYLEDADLALAKNIYAGAFYSGORCTAFKRVLYVMDKVAADL 300
Qy 301 AEKISENVAKLSVGDPPDNATVTPVDDNSADIEELVYDARQKAEINEFGRDRLLT 360
Db 301 AEIKTLVKLSVGMPEDDADITPLDTSADFBEELIDARDKGTALTAFFRREGNLLS 360
Qy 361 PGLFDHVTLLDMKLAEEEPGPIPIIRVKDAEEAVALANKSDFGLOSSVFTDPOKAFDI 420
Db 361 PVLFDHVTLLDMKLAEEEPGPVPIIRVTVVEALIKISNEBSYGLQASLFTTNFPAFPI 420
Qy 421 ANKLEVGVTHINNKTRGPDNPPLGLKSGAGVQIGIRYSIAMTVKSIIVLDMK 475
Db 421 AEGLEVGVTHINNKTRGTDNFPPLGAKKSGAGVQVKSIAMTVKSIIVFDIQ 475

RESULT 8
US-10-282-122A-53543
Sequence 53543, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931

```
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 53543
/ LENGTH: 486
/ TYPE: prt
/ ORGANISM: Clostridium difficile
US-10-282-122A-53543
```

```
Query Match          58.4%; Score 1397; DB 15; Length 486;
Best Local Similarity 56.9%; Pred. No. 2e-103;
Matches 272; Conservative 80; Mismatches 122; Indels 4; Gaps 2;
```

```
QY 2 TTE--YQNVNGEW--KSSVNOIEIISPIDSSLGFPVPMTRREVDHAKAGREALPAMA 57
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 7 TKENYKNLNGKMWESNRKPIEIVSPIDNSLVGVQSGTKHEVDVKNTKESIKVMA 66
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 58 ALTVERAQLAKADIERDKEEATVLAKEISKAYNASVTEVATDLIRYAABEGIR 117
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 67 EMPVTKRANIFPKAADLLENIIDELANILVLEIAKDIKARAEVERTADFLRTADVGN 126
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 118 ISTSDDEGKMDASTGSKLAVIRPQVGIVLAIPYNPVNLGSKIADPALIGNVVMFK 177
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 127 MEGEALISGNFPGGRTRNKNYSVSRVPLGIVLAISPNYVNLMSKIAIALIGNANVLAK 186
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 178 PPTQSSVSGVLAKAPAEAGLPAGVNTITGSGSEIGDIYVHEEVNPFNFTGSPVQGR 237
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 187 PATQGISALHAYEIRKAGIPDGVNLTITGSGSEIGDIYVTHKGINFTGSTEVGQH 246
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 238 ICKLAGMRPIMLEIGSKDGIYLAADADLNAKOIYAGAYDSGGCTAIKRVLYVEEA 297
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 247 IKSISMVNLLLELGGKDAIVLEDDADLPAAKNIVSGAYSGQCTAVKILVQESYA 306
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 298 DELAEKISENVAKLTVGDPFNATVTPVIDNSADPIESLVVDARQKAKELNEFRDGR 357
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 307 DLTVGKIKRVLKLTIGNMDEVITPLIDNKATDPVQGLVDDALHKAKLITGNVRKN 366
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 358 LITPLGFHVTLDMLKAWEEPFGPIIPITRVKDAEBAVAIAKNSDFGLQSSVFTDFOKA 417
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 367 LVEYFTLLDVANVDMKIAWEEFPGVLPITIRVDINQAIHIANQSEVGLQSSVFTSDIDKA 426
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 418 FPIANKLEVTGHINKKTRGPDNPFPLGLKSGAGVQGIKRSIEMTVKSVIIVDMK 475
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 427 FYIADKLEVTQVQINNKTGPDHFPFLDVKASGMGTQGVKYSIEMTRPKAVVNVNR 484
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

RESULT 9

```
US-10-282-122A-52739
/ Sequence 52739, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA 034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
```

```
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 52739
/ LENGTH: 486
/ TYPE: prt
/ ORGANISM: Clostridium botulinum
US-10-282-122A-52739
```

```
Query Match          58.0%; Score 1388.5; DB 15; Length 486;
Best Local Similarity 56.7%; Pred. No. 9.9e-103;
Matches 268; Conservative 90; Mismatches 112; Indels 3; Gaps 2;
```

```
QY 5 YQNVNGEWSV--NQIEIISPIDSSLGFPVPMTRREVDHAKAGREALPAMAALTVY 62
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 12 YKNFPGQWNSKTNKPFIEISSPIDSGLVGVPSKKEVDLAIQAKAGSHMNEIRLIN 71
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 63 EBAQYLKRAADIERDKEEATVLAKEISKAYNASVTEVATDLIRYAABEGIRLSTSA 122
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 72 EKATILKAAANILDEKAEIADIMTKETAKOKSSISEVRRADYIRFSADTAKMVGET 131
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 123 DEGGKMDASTGSKLAVIRPQVGIVLAIPYNPVNLGSKIAPLIGNANVMPFPQOG 182
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 132 LPGDSFPGNSKSKSIVKRVPLGVVLAISPNYVNLGSKIAPLMAKNSVVLKPAQOG 191
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 183 SVSGVLAKAPAEAGLPAGVNTITGSGSEIGDIYVHEEVNPFNFTGSPVQGRIGKLA 242
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 192 SISALHAKVFEBAQIPSGVNLITGSGSEIGDIYVSPMIDFINFTGSTEVGKRISSH 251
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 243 GMRPIMLEIGSKDGIYLAADADLNAKOIYAGAYDSGGCTAIKRVLYVEEVDLAE 302
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 252 VMKPMLEIGSKDAIVLEDDADLPAAKNIVSGAYSGQCTAVKILVLEIADTLVE 311
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 303 KISENVAKLTVGDPFNATVTPVIDNSADPIESLVVDARQKAKELNEFRDRLITP 361
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 312 KYKERVESLKTGNPKQVNDIVPLIDKAAAYVEALMEBANKKATLLVGGREBNLYP 371
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 362 GLFDTVLDMLKAWEEPFGPIIPITRVKDAEBAVAIAKNSDFGLQSSVFTDFOKAFDIA 421
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 372 TLFEDVTLDMLKAWEEPFGPIIPITRVKDVDEALIAHESERGLOASLFTKEINKAFVYA 431
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 422 NKLAVGVTHINKKTRGPDNPFPLGLKSGAGVQGIKRSIEMTVKSVIIVDMK 474
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 432 EKLAVGVTVQINNKPGRPDHFPFSGIKASGLGTQGIKYSIEMSRPKAIVINI 484
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

RESULT 10

```
US-10-282-122A-52073
/ Sequence 52073, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
```


177 KPPTGSGVGLVAKAFAGLPAQVFNITGSGSRIGDVIYHEHVNFINFGSPVQ 236
176 KSTQSSGLIGIRISELVHEVGPFGVAVNCLITTEARITGDLVTKVKAISFGGPKVN 235
237 RIGLKAMPIMLEGGDAGIYVADADLDNNAKQIVAGAYDSGORTAIRVLVEEV 296
236 HIAETSKISLVLEGGKOPALVLDADPELANELVKAIGYSGGRCRAIKRVFVSHN 295
297 ADELAKISENVAKISVGPDPNATVPIYIDNSADFISSLVDAROKA--KELEFR 354
296 HDLVNVLINKKVDALVGLPQGNPITPLINSNLSLYNLSLVEDAIKKAIVHOKVYNE 355
355 DGRILTPGFDVYTLDMKLAWEPRGPILPIRVDAEAVAMANKSDGLOSVFTRDF 414
356 KNNLHPLVIDNTKEMRVAMEPFGPILPIITYNSIQEALIDLINSOYQLOACTYVY 415
415 OKAFDIANKLEVGTVINNKGTGPNPPLGLKSGAGVQIGRYSIEAMTNKSVLD 473
416 ASIEQALQIESGTININKSSSRGPILPFGVKDSGFGVQIVDAISMTTKGIIN 474

RESULT 19
US-10-268-518-4
; Sequence 4, Application US/10268518
; Publication No. US20030100034A1
; GENERAL INFORMATION:
; APPLICANT: Hunter, John Joseph
; TITLE OF INVENTION: 9136, A HUMAN ALDEHYDE DEHYDROGENASE
; FILE REFERENCE: MP101-234PIRM
; CURRENT APPLICATION NUMBER: US/10/268,518
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/329,899
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 492
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
US-10-268-518-4

Query Match 32.9%; Score 786.5; DB 14; Length 492;
Best Local Similarity 38.6%; Pred. No. 2.3e-54;
Matches 193; Conservative 92; Mismatches 166; Indels 49; Gaps 16;

12 EWKSSVN--QIEILSPIDSS--IGFVPAMTREVDHAKAGREAL--PAMALTYERA 65
1 EWVDSASGTFEEVAVPANKGEVIGRVPEATADVDAAVAAKAFKSGPMWAKVPASEBA 60
66 QVLHKAADIIEERKEIATVLAKEISK--AYNASVTEVVRTADLRYYA-----EBG 115
61 RILRKADILIEERDELALETLDLGPPLAEAKGDTVEGALDEIRYVAGMARKMGERR 120
116 I--RISTSADEGGKMDASTGHKLAVIRPOVGIIVLAIRYVYVNLGSKIAIPALIGNV 173
121 VLPISLATDDE-----ELNTRYRREPLGVGVISPMNPFLLALMKLAPALAAQNT 170
174 VNFKEPTQGSVSGVLAKAFPAEG--LPAGVFNTITGSGSEIGDVIYHEHVNFINFG 230
171 VVLPKPEEQPLTALMLAEILIEEGANLTPKGVVAVVPGCAEVGQALLSHPDIDKISFG 230
231 STPVGGRIGKLA---GMRPIMLELGGKDGIVLADADLDNNAKQIVAGAYDSGORTAI 287
231 STEVGKILIEAARAAKUKVLTTELGGKSPYIVFDDDLDAVRIYFGAFGNAQVCYIAP 290
288 KRVLYVEEVADELAEKISENVAKLS--VGDPFNAT--VTPVIDNSADFISSLVDAROK 344
291 SRLVHESIYDEVEVKERVKTKLIGPLDSDTNIGYPLISEQGFDRVLASYIEDGKEB 350
345 GAK-----ELNFRKRGGRLLTGLFPHVTLDMKLAWEPRGPILPIIRVDAEAVYA 396

351 GAKVLGGERBDSKELYLGQYVQPTIFTDVTPDKINKKEILFGVLPILRKDLDRAIE 410
397 IANKSPFGLQSSVFTRD--FQKAFDIANKLEVGTVINNKGTGPD--NPPFLGK--GSGAG 453
411 LANDTEYGLAAVVPFKDILARAFVAKALEAGIVVAVNDVCVAABPOLPFGGVKQSSGIG 470
454 VO-GIRYSIEAMTNKSVTL 472
471 REHGKGYGLEEYTEIKTVTI 490

RESULT 20
US-10-391-364-28
; Sequence 28, Application US/10391364
; Publication No. US20040121349A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Cook, William James
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Welch, Nadine S.
; APPLICANT: Bandaru, Rajasekhhar
; TITLE OF INVENTION: NOVEL 27877, 18080, 14081, 32140, 50352,
; TITLE OF INVENTION: 16656, 14223, 16002, 50566, 65552 AND 65577 MOLECULES AND
; FILE REFERENCE: MP103-019QNMIM
; CURRENT APPLICATION NUMBER: US/10/391,364
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/950,370
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/231,084
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/338,587
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 10/266,035
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US 60/328,198
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 09/717,926
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: US 10/268,036
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/327,820
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/327,820
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 492
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Aldehyde dehydrogenase Pfam consensus sequence
US-10-391-364-28

Query Match 32.9%; Score 786.5; DB 16; Length 492;
Best Local Similarity 38.6%; Pred. No. 2.3e-54;
Matches 193; Conservative 92; Mismatches 166; Indels 49; Gaps 16;

12 EWKSSVN--QIEILSPIDSS--IGFVPAMTREVDHAKAGREAL--PAMALTYERA 65
1 EWVDSASGTFEEVAVPANKGEVIGRVPEATADVDAAVAAKAFKSGPMWAKVPASEBA 60
66 QVLHKAADIIEERKEIATVLAKEISK--AYNASVTEVVRTADLRYYA-----EBG 115
61 RILRKADILIEERDELALETLDLGPPLAEAKGDTVEGALDEIRYVAGMARKMGERR 120

QY 116 I--RLSTADDEGKMDASTGTHKLAVIRQPVGI VLAIAFYNTPVNLGSGKIAPALIGNV 173
DB 121 VLPISLATDDE-----ELNTRYRREPLGVGVISPMNFPLLALWKLPALAAAGNT 170
QY 174 VNFKEPPTQSSVGLVLAKAFAEAG--LPAQVNTITGSGSEIGDYVHEBEVNFINFG 230
DB 171 VVLKSEQPTPLPALLAELIEEGANNLKGVVNVPGFAEGQQLLSHPIDIKISFPG 230
QY 231 STPVGQRIGKLA---GMRPIMLELGKADAGI VLAIDLNAKAQIYAGADYSGQRCTAI 287
DB 231 STEVGLKLIWEAAAANKLKVITLELGGKSPVIVPDDDLDKAVERIVFGAFNAGQVCIAP 290
QY 288 KEVLVEEVADELAEKISSENVAKLS-VGDPFDNAT--VTPVIDNSADPIESLVDAROK 344
DB 291 SRLVHESIYDEFEVKEKERVKKLIGDPLDSDTNIYGPLISEQGFDRVLSTIEDGKEE 350
QY 345 GAK-----ELNEFRKDRLLTPGLFDHVTLLDKLWEEPPGPIILPIIRVKDAEENAV 396
DB 351 GAKVLGGERDSKREYLGSGYVQPTIFTDVTPDMKIMEKEIFGPVLPILIKERDLEALE 410
QY 397 IANKSDPGLQSSVFTRD-FOKAFDIANKLEVGTVHINNTGGRPD-NPPLGLK-GSGAG 453
DB 411 LANDTEYGLAAVFTFDILARAFAVKALEAGI VWNVDVCVAHAEPQLPFGGVKQSSGIG 470
QY 454 VO-GIRYSIEMNTNKSIVL 472
DB 471 REHGKYGLEBYEYIKVTI 490

RESULT 21

US-09-823-901-9
/ Sequence 9, Application US/09823901
/ Patent No. US20020001807A1
/ GENERAL INFORMATION:
/ APPLICANT: Meyers, Rachel
/ TITLE OF INVENTION: 21509 AND 33770, NOVEL HUMAN
/ FILE REFERENCE: 10448-036001
/ CURRENT APPLICATION NUMBER: US/09/823,901
/ CURRENT FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: US 60/193,920
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 9
/ LENGTH: 493
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: consensus sequence
US-09-823-901-9

Query Match 32.5%; Score 777; DB 9; Length 493;

Best Local Similarity 38.5%; Pred. No. 1.3e-53; Indels 50; Gaps 17;

Matches 193; Conservative 92; Mismatches 166; Indels 50; Gaps 17;

QY 12 EKKSSVN--QIEILSPIDSS--LGFVPAMTREBEVDHAKAGREAL---PAMALTVYERA 65
DB 1 EAVDASAGKTFEVNPNANGEVIGRPEATIEDVDAAVAAKAEAFSGGMWAKVPASERA 60
QY 66 QYLAHRAADIIEERKEIATVLAKEISK--AYNASVTEVVRTADLIRYAA-----BEG 115
DB 61 RLRLKLAADIIEERDELTALLETLDLCKPLAEAKGDTREVRADIEIRYVGMARKLMGERR 120
QY 116 I--RLSTADDEGKMDASTGTHKLAVIRQPVGI VLAIAFYNTPVNLGSGKIAPALIGNV 173
DB 121 VLPISLATDDE-----ELNTRYRREPLGVGVISPMNFPLLALWKLPALAAAGNT 170
QY 174 VNFKEPPTQSSVGLVLAKAFAEAG--LPAQVNTITGSGSEIGDYVHEBEVNFINFG 230
DB 171 VVLKSEQPTPLPALLAELIEEGANNLKGVVNVPGFAEGQQLLSHPIDIKISFPG 230
QY 231 STPVGQRIGKLA---GMRPIMLELGKADAGI VLAIDLNAKAQIYAGADYSGQRCTAI 287

DB 221 STEVGLKLIWEAAAANKLKVITLELGGKSPVIVPDDDLDKAVERIVFGAFNAGQVCIAP 290
QY 288 KEVLVEEVADELAEKISSENVAKLS-VGDPFDNAT--VTPVIDNSADPIESLVDAROK 343
DB 291 SRLVHESIYDEFEVKEKERVKKLIGDPLDSDTNIYGPLISEQGFDRVLSTIEDGKEE 350
QY 344 GKAK-----ELNEFRKDRLLTPGLFDHVTLLDKLWEEPPGPIILPIIRVKDAEENAV 395
DB 351 EGAKVLGGERDSKREYLGSGYVQPTIFTDVTPDMKIMEKEIRGPVLPILIKERDLEAI 410
QY 396 AIANKSDPGLQSSVFTRD-FOKAFDIANKLEVGTVHINNTGGRPD-NPPLGLK-GSGA 452
DB 411 ELANDTEYGLAAVFTTKIILARAFAVKALEAGI VWNVDVCVAHAEPQLPFGGVKQSSGI 470
QY 453 GVO-GIRYSIEMNTNKSIVL 472
DB 471 REHGKYGLEBYEYIKVTI 491

RESULT 22

US-10-175-696-21
/ Sequence 21, Application US/10175696
/ Publication No. US20030092658A1
/ GENERAL INFORMATION:
/ APPLICANT: Glucksmann, Maria Alexandra
/ APPLICANT: Meyers, Rachel
/ APPLICANT: Rudolph-Owen, Laura A.
/ TITLE OF INVENTION: NOVEL HUMAN ENZYME FAMILY MEMBERS AND USES THEREOF
/ FILE REFERENCE: 10448-193001
/ CURRENT APPLICATION NUMBER: US/10/175,696
/ CURRENT FILING DATE: 2002-06-20
/ PRIOR APPLICATION NUMBER: 10/067,668
/ PRIOR FILING DATE: 2002-02-04
/ PRIOR APPLICATION NUMBER: 60/266,140
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 09/823,901
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: PCT/US01/10720
/ PRIOR FILING DATE: 2001-04-02
/ PRIOR APPLICATION NUMBER: 60/193,920
/ PRIOR FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 09/862,658
/ PRIOR FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: PCT/US01/16380
/ PRIOR FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: 60/205,675
/ PRIOR FILING DATE: 2000-05-19
/ PRIOR APPLICATION NUMBER: 09/882,837
/ PRIOR FILING DATE: 2001-06-15
/ PRIOR APPLICATION NUMBER: PCT/US01/19319
/ PRIOR FILING DATE: 2001-06-15
/ PRIOR APPLICATION NUMBER: 60/211,727
/ PRIOR FILING DATE: 2000-06-15
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 21
/ LENGTH: 493
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: consensus sequence
US-10-175-696-21

Query Match 32.5%; Score 777; DB 14; Length 493;

Best Local Similarity 38.5%; Pred. No. 1.3e-53; Indels 50; Gaps 17;

Matches 193; Conservative 92; Mismatches 166; Indels 50; Gaps 17;

QY 12 EKKSSVN--QIEILSPIDSS--LGFVPAMTREBEVDHAKAGREAL---PAMALTVYERA 65
DB 1 EAVDASAGKTFEVNPNANGEVIGRPEATIEDVDAAVAAKAEAFSGGMWAKVPASERA 60
QY 66 QYLAHRAADIIEERKEIATVLAKEISK--AYNASVTEVVRTADLIRYAA-----BEG 115


```

Qy 12 EKKSVN--QIILSIDSS--LGFVPAMTREEDVHAMKAGREAL---PAMAAITVERA 65
    1 EWVDSASGKTFEEVNVANKGEVIGRVPEATAEDVDAVAAKAEAFKSGPMWAKVPSERA 60
Db 66 QYIHRADIIERDKKEIATVLAKEISK--AYNASTEYVTRADLIYYA-----SEG 115
    61 RILRIKADIIEEBEDLALLETDLCKPLAEAKGTEVERALDEIRYYGAMARKMGERR 120
Qy 116 I--RLSTSDGKMDASTGHLKLAIVRQPVGIVLAIPYNPVNLSGSKIAPALIGNV 173
    121 VPLSLATDDE-----ELNTRYRREPLGVGVGISPMNPFLLALMKLAPALAGNT 170
Db 174 VMEKPEPTGSGVGLVLAFAEAG--LPAGVFNTITGSGSEIGDYVHEEYVNFINTG 230
    171 VVLKSEQTPPLTALLAEILIEEGANNLPKGVNVNVPFGAEVQALLSHPIDIKISFTG 230
Qy 231 SPPVQRIKLA---GMRPIMLELGKDKAGIYLAADLDNNAKOIYAGAYDSGQRTAI 287
    231 STEVGLIMEAAAKVLLKVTLELGKSPVIYFDDLDLKAVERIYFAGFAGNAGQVCIA 290
Qy 288 KEVLVVEEVADELAEKISENVAKL-S-VGDPFDNAT--VTPVIDNSADPTESLVVDAROK 344
    291 SKLLVHESIYDEFEVEGLKRVKRLKLGIDPLDSDTNIYGPLISEQOPDRLVLTIEDGKE 350
Qy 345 GAK-----ELNEFKRDLGRLTLPGLFDHVTLDMLAMEEPPGPIIPIRVDAEAVA 396
    351 GAKVLCGGERDSKEVLCGGYVQPTIFDVPDMKIMEELFGVPLIIFKFDLDEALE 410
Qy 397 IANKSDFGIQQSVFTPD-FOKAFDINKLEVGTVHINNKTGRGPD-NFFPLGL-KSSGAG 453
    411 LANDVEYGLAAVYFTDILARAFVRKALEAGIWMVDVCVAHAEPQLPFGVHQSIGIG 470
Qy 454 VO-GIRYSIEAMTNVKSIVL 472
Db 471 REHGKGYGLEYTEIKTVI 490

RESULT 25
US-10-369-493-23067
; Sequence 23067, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23067
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-369-493-23067

Query Match      31.9%; Score 763; DB 15; Length 488;
Best Local Similarity 36.9%; Pred. No. 1,7e-52;
Matches 170; Conservative 94; Mismatches 179; Indels 18; Gaps 6;

```

```

Qy 122 ADEGKMDASTGHLKLAIVRQPVGIVLAIPYNPVNLSGSKIAPALIGNVMEKPEPTQ 181
    128 ---GDVISTDQDALMPTTRVPLGVGVGISPMNPFVAILPKWMAALVYGMTVVIKPAE 184
Qy 182 GSVGLVLAKAPAEAGLPAQVFNITGSGSEIGDYVHEEYVNFINTGSPVQRIKGL 241
    185 TAVTCAKIIACEEELPAGVINYLVITGSGSVVGGGLAHGDVNAATFSGNQVGIIOQA 244
Qy 242 AGMR--PIMLELGKDKAGIYLAADLDNNAKOIYAGAYDSGQRTALKRVLVVEVADE 299
    245 ALARQAKIQLKEGKNPVIYADDLLENAAEVITGARSTRGQCTATSRVIVSGIYER 304
Qy 300 LAEKISENVAKL-SVGDPE-DNATVTPVIDNSADPTESLVVDARQAKEL-----NE 351
    305 FKEKLLQRTKQITTDLSKEDEVMWGPISAKQDLWCLSYIEKGQEGASLLIGKEKENG 364
Qy 362 FKRDRLTLPGLFDHVTLDMLAMEEPPGPIIPIRVDAEAVAIANKSDPGLQSSVFT 411
    365 KYQNGYVQPAIFDVTSEMTIAQEIPIGPIYALIKVDSIEBALNIANDVKGSLASIFT 424
Qy 412 RFOKAFDINKLEVGTVHINNKTGRGPDNFFPLGLKSGA 452
Db 425 ENIGMSPFIDEIDAGLVINAESAGVELQAPFGMKQSS 465

RESULT 26
US-10-369-493-1073
; Sequence 1073, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1073
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-10-369-493-1073

Query Match      31.7%; Score 759.5; DB 15; Length 463;
Best Local Similarity 37.1%; Pred. No. 3.1e-52;
Matches 175; Conservative 95; Mismatches 183; Indels 19; Gaps 8;

```

```

Qy 5 YQNVNGEW-KSSVNOIILSPID-DSSLGFVPAMTREEDVHAMKAGREALPAMAAITV 61
    10 YNFIIGEEVWKSQSGMNVAVENDADVNDIVGVONSTADVERAVYAANEAKTAMKRLNG 69
Db 62 YERAOVLHKAADIIEERDKKEIATVLAKEISKAYNASVTEVVRADLIRVAAEGRILSTS 121
    70 AERGOVLTYKTADIEORLEIIAACATREMGKTLPEAKGETAGIALIRYVAGEGMKMT-- 127

```

```

Qy 122 ADEGKMDASTGHLKLAIVRQPVGIVLAIPYNPVNLSGSKIAPALIGNVMEKPEPTQ 181
    128 ---GDVISTDQDALMPTTRVPLGVGVGISPMNPFVAILPKWMAALVYGMTVVIKPAE 184
Qy 182 GSVGLVLAKAPAEAGLPAQVFNITGSGSEIGDYVHEEYVNFINTGSPVQRIKGL 241
    185 TAVTCAKIIACEEELPAGVINYLVITGSGSVVGGGLAHGDVNAATFSGNQVGIIOQA 244
Qy 242 AGMR--PIMLELGKDKAGIYLAADLDNNAKOIYAGAYDSGQRTALKRVLVVEVADE 299
    245 ALARQAKIQLKEGKNPVIYADDLLENAAEVITGARSTRGQCTATSRVIVSGIYER 304
Qy 300 LAEKISENVAKL-SVGDPE-DNATVTPVIDNSADPTESLVVDARQAKEL-----NE 351
    305 FKEKLLQRTKQITTDLSKEDEVMWGPISAKQDLWCLSYIEKGQEGASLLIGKEKENG 364
Qy 362 FKRDRLTLPGLFDHVTLDMLAMEEPPGPIIPIRVDAEAVAIANKSDPGLQSSVFT 411
    365 KYQNGYVQPAIFDVTSEMTIAQEIPIGPIYALIKVDSIEBALNIANDVKGSLASIFT 424
Qy 412 RFOKAFDINKLEVGTVHINNKTGRGPDNFFPLGLKSGA 452
Db 425 ENIGMSPFIDEIDAGLVINAESAGVELQAPFGMKQSS 465

RESULT 26
US-10-369-493-1073
; Sequence 1073, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1073
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-10-369-493-1073

Query Match      31.7%; Score 759.5; DB 15; Length 463;
Best Local Similarity 37.1%; Pred. No. 3.1e-52;
Matches 175; Conservative 95; Mismatches 183; Indels 19; Gaps 8;

```

QY 302 EKISENVAKLSVDPDFDNAT-VTPVIDNSADFIESTLVVDAROKAKELNEFKRDGLLT 360
 Db 292 EMFVNNAKVLNVNPNDEKTDVGPILISVEHAWEVEKVEALDEGKLLGGKRDALFY 351
 QY 361 PGLFDHVTLLDMKLAMEEPPGPIIPIIRVKDAEEAVALANKSDFGLOSVFTDFOKAFDI 420
 Db 352 FTLE-VDRDNILCKETETFAVPIPIRTWE-EEIMDIANSTEGYLSAIFTNDINKSLKF 409
 QY 421 ANKLEVTGHIINNKTRGPNPFLGLKSGAGVOCIRYSIEMTVKSVYL 472
 Db 410 AENLEFGVAVINDSSLPRDNNPFGVKSGLREGVXKAMEEBSNIKTIII 461

RESULT 27

US-10-369-493-1188
 ; Sequence 1188, Application US/10369493
 ; Publication No. US2003023675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 1188
 ; LENGTH: 455
 ; TYPE: PRT
 ; ORGANISM: Mechanobacterium thermoautotrophicum
 US-10-369-493-1188

Query Match 31.0%; Score 741.5; DB 15; Length 455;
 Best Local Similarity 38.2%; Pred. No. 8,4e-51;

Matches 173; Conservative 83; Mismatches 186; Indels 11; Gaps 5;

QY 9 VNGEMSSVNOQIILSPIDSSLGFPVPMATREEDVHAMKAGREALPAMMALTYERAQYL 68
 Db 5 IDDEGVSGERKIFVRNRPNGDEVDVPLAGRSVEYEAIRAHRAAMADLSARKISEL 64
 QY 69 HKAAIIERDKKEIATVLAKESIKAYNASVTEVVRTADLIRYAEGBIRLSTSADEGKM 128
 Db 65 YVDADELKTELDEFARLITLESKPIRFSRDEYKRSVETARLSAEAGRL--YGESIM 121
 QY 129 DASTGK--LAVIRROPVGIVALAIPYVNVNLSSKIAIPALIGNVVMFKPTQSVG 186
 Db 122 DAGIGGKGLTFTVRIPLGVAALITFPNYPNLAIHKVGPALAAAGTTILKSLLEAPLSA 181
 QY 187 LVYAKFAEAGLPAGVFNITGSGSEIGDYIVEHEEVNFINTGSTPGORIGKLAGMP 246
 Db 182 LKLAWLSE-HPPAGAVNAVYGRGSEVGVIIIDSPVDTKITFTGSEVGVISARASMK 240
 QY 247 IMLELGKAGIYVLADADDNNAKOIVAGAYDYSQGRCTAIKRVLVVEEVADELAEKISE 306
 Db 241 ITLELGGNDPLIYMDADIDSAVEAAVRSYLSGVCTIAKRMIVHEVDADDEFADKLVN 300
 QY 307 NNAKLSVGPFPDNAT-VTPYIDNSADFIESTLVVDAROKAKELNEFKRDGLLTGFLD 365
 Db 301 ITGSLAAGDPMVDRITVGPLINDDAIIEVERVIGAIVEGAELCCGSRGNFVEPTVAD 360
 QY 366 HTYLDKLAWEPPGPIIPIIRVKDAEEAVALANKSDFGLOSVFTDFOKAFDIANKKE 425
 Db 361 HAVPGMEVIERETFGVSPRIIRNGADEAIRINGCYALQAGVTENIRITALMAKEIE 420
 QY 426 VGTVHINNKTRGPNPFLGLKSGAGVOCIR 458
 Db 421 AGTVLVNKQSTFRVDHMPF---GDSGAVEWVR 449

RESULT 28

US-09-738-626-6430
 ; Sequence 6430, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 6430
 ; LENGTH: 484
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-6430

Query Match 31.0%; Score 741.5; DB 9; Length 484;
 Best Local Similarity 36.9%; Pred. No. 9,2e-51;

Matches 174; Conservative 93; Mismatches 184; Indels 21; Gaps 6;

QY 6 QNYVNGEW-----KSSVNOQIILSPIDSSLGFPVPMATREEDVHAMKAGREALPAMMAL 59
 Db 12 KHLIGQWEGNSDIDSTV---INPYDSVIAESKQSIADVDAAVEAKKAQKQEMMAT 67
 QY 60 TYERAQYILHKAADIIERDKKEIATVLAKESIKAYNASVTEVVRTADLIRYAEGBIRLS 119
 Db 68 PAERSAIIYRAAELEBEHREIIVEMLIKESGSTSKANLEITLAGNTKESASPPGRV- 126
 QY 120 TSADGKMKDASTGKLAIVIRROPVGIVALAIPYVNVNLSSKIAIPALIGNVVMFKPP 179
 Db 127 ----HGRISPSTPBEKENVYRVAGGVGVISPMNPLNLSIRSVAPALAVGNAVITKPA 182
 QY 180 TQGSVSGVLV-AKAPAEAGLPAGVFNITGSGSEIGDYIVEHEEVNFINTGSTPGORI 238
 Db 183 SDTPVTGVIPARIFEEAGVPAGVISTVAGAGSEIGDHVTVAVPKLISFTGSTPGRRV 242
 QY 239 GYLA-----GMRIIMLELGKAGIYVLADADDNNAKOIVAGAYDYSQGRCTAIKRVLYVE 294
 Db 243 GELAINGPMKTVALELIGNAPFVVLADADIDAAQAAVGAFLHQGOICMSINRYIYDA 302
 QY 295 EYADELAEKISENVAKLSVDPDFDNAT-VTPYIDNSADFIESTLVVDAROKAKELNEFK 353
 Db 303 AVHDFLEKFEVAVKNITPRGDSAGCTIVGPIINDSGLGKEKIELAKKEATVQVEGP 362
 QY 354 RDGRLITPGLFDHVTLLDMKLAMEEPPGPIIPIIRVKDAEEAVALANKSDFGLOSVFTFD 413
 Db 363 IEGRLVHPHFSDVTSDEIAEERIFGPIISYLVKADDDRAHAELANASDFGLSAVWSKD 422
 QY 414 FOKAFDIANKKEVGTVHINNKTRGPNPFLGLKSGAGVOCIRYSIEMTVKSVYL 465
 Db 423 IDRAAQFALQIDSGVHINDLTVNDEPHVMFGSKNSGLGRPNGDWALIEEFT 474

RESULT 29

US-10-781-014-228

```
/ Sequence 228, Application US/10781014
/ Publication No.: US20040180408A1
/ GENERAL INFORMATION:
/ APPLICANT: Pompejus, Markus
/ APPLICANT: Kroeger, Burkhard
/ APPLICANT: Schroder, Hartwig
/ APPLICANT: Zelder, Oskar
/ APPLICANT: Haberer, Gregor
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
/ TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
/ TITLE OF INVENTION: PRODUCTION
/ FILE REFERENCE: BGI-126PCN
/ CURRENT APPLICATION NUMBER: US/10/781,014
/ CURRENT FILING DATE: 2004-02-17
/ PRIOR APPLICATION NUMBER: US 09/602,740
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/141,031
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: 60/143,208
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: 60/151,572
/ PRIOR FILING DATE: 1999-08-31
/ PRIOR APPLICATION NUMBER: DE 19931412.8
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931413.6
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931419.5
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931420.9
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931424.1
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931428.4
/ PRIOR FILING DATE: 1999-07-08
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 784
/ SEQ ID NO 228
/ LENGTH: 484
/ TYPE: PRF
/ ORGANISM: Corynebacterium glutamicum
US-10-781-014-228

Query Match          31.0%; Score 741.5; DB 16; Length 484;
Best Local Similarity 36.9%; Pred. No. 9.2e-51;
Matches 174; Conservative 93; Mismatches 184; Indels 21; Gaps 6;

QY 6 QYVYNGEM-----KSNVQIETLSPIIDSSLGFPVPMTRREVDHMKAGRELPMAALTY 59
DB 12 KHLGGQWEGNSDRISTN---INPYDSDVIAESKQASIDADVAEAKKAQAEWMAAT 67
QY 60 TVYERAOYLHKAADIIRERKEEIAATYLAKEISKAVNASVTEVVRTADLIRYAEERGIRLS 119
DB 68 PAERERATITRYRAHELEERREELVEWLIRESGSTSKANLEITLAGNITKESHPFGRV- 126
QY 120 TSADEGKMDASTGHLKAVIRROPVGIYLAIPYNPVNLSGSKIAPALIGNVVMFKRP 179
DB 127 ----HGRISPSTPKGENRYVYRAKGVGVISPMNPFNLSTIRSVAPALAVGAAVITKDA 182
QY 180 TGGSVSGVL- AKAPAEAGIPAGVFNITIGRSGEIDYIVHEBEVNFNFTGSTPVGQRI 238
DB 183 SOTPVYGVIPARIFEEAGVPAGVISITVAGAGSEIDHVTAVVPLKLIPTSTSPGRGV 242
QY 239 GKLA-----GMRPIMLELGKDGIVLADADLDNAKOIYVAGAVDYGSGRCTAIKRYLYVE 294
DB 243 GELATNGEMKTYALLEGNAFPVVLADADIDAAQAALVAGATLHGQOICMSINRYIVDA 302
QY 295 EVADELAERKISENVAKLSVGDPEPDNAT-VTPVIDNSADFIETSLVVDAROKAKELNEBK 353
DB 303 AVHDEFLERKEVEAVKNIPTGDPFABEGTLVGPVINDSQSLGKKEIKELAKKEGATVQVEGP 362
QY 354 RGRLLTPTGLFDHYTLIDMKLAMEEPFGPIIPIIRYVDAEBAVALANKSPFGLOSSFPTTD 413
DB 363 IEGRLVHPVFSVDSYDMEIAREEIPGLISVLKADDEAHAAETLANASDFGLSAAVMSKD 422
```

```
QY 414 FOKAFDIANKLEVGTVHINNKTGRGPDNPFGLKSGAGVQGIKRYSEIANT 465
DB 423 IDRAAQFALQIDSGVNHINDLTVDNDEPHVMGSGKNSQLGFNDMAIEEFT 474

RESULT 30
US-10-369-493-17
/ Sequence 17, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 17
/ LENGTH: 476
/ TYPE: PRF
/ ORGANISM: Aquifex aeolicus
US-10-369-493-17
```

```
Query Match          30.8%; Score 737; DB 15; Length 476;
Best Local Similarity 35.0%; Pred. No. 2.1e-50;
Matches 168; Conservative 96; Mismatches 200; Indels 16; Gaps 6;
```

```
QY 4 EYQYVNGEMKSVNQIETLSPIIDSSLGFPVPMTRREVDHMKAGRELPMAALTY 63
DB 3 EKKMLIGGEWWDKEEIVIIPYTRKPIGRVPKDEKVERAIEBAKKEFEISLTAYE 62
QY 64 PAQYHKAADIIRERKEEIAATYLAKEISKAVNASVTEVVRTADLIRYAEERGIRLS--- 119
DB 63 RYELMRANQOLKEBAEFPAKTLVLEVGKTRREARTEVQRAIQTLIFSAEAKVNGETF 122
QY 120 -TSADEGKMDASTGHLKAVIRROPVGIYLAIPYNPVNLSGSKIAPALIGNVVMFKRP 178
DB 123 PIDAHPNCK-----GKLGFIYRVPGIVSAITPFPNLMSMKHVAADLAAGNAVILKP 176
QY 179 PTQVSGVLAKAPAEAGIPAGVFNITIGRSGEIDYIVHEBEVNFNFTGSTPVGQRI 238
DB 177 SERTPLTPLMGEILLEAGVPPKALSVIPGVG-DVGKAMTTHPPDVRYVSPFGSRKRVGDII 235
QY 239 GKLAEMRPIMLELGKDGIVLADADLDNAKOIYVAGAVDYGSGRCTAIKRYLYVEEYAD 298
DB 236 ARQVGIKCLVLELGSNSAIIILHKGDNLEKAVETVQGYAALAGQVCISVQRFVHEDIYD 295
QY 299 ELAEKISENVAKLSVGDPEPDNAT-VTPVIDNSADFIETSLVVDAROKAK-ELNEFK--R 354
DB 296 EITKRLKERVSLKLVGDPMNEDTDLGPMIARSEIERIGEWISBAVQKAKIEAGGLRCAB 355
QY 355 DGRLLTPTGLFDHYTLIDMKLAMEEPFGPIIPIIRYVDAEBAVALANKSPFGLOSSFPTD 414
DB 356 DETTSPPTIVSLVPPDSKLFKEEAPAPVAVANPNYDIEEALRWNSSSYGLOVGFITDV 415
QY 415 OKAFPIANKLEVGTVHINNKTGRGPDNPFGLKSGAGVQGIKRYSEIANTVKSIVLDM 474
DB 416 KLAHCITEIEAGVVLINBGPFRADHNPYGCVKSGIGREGPKAIEDYIEIKTVITDL 475

RESULT 31
US-10-369-493-17249
/ Sequence 17249, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
```



```

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17249
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Bacillus halodurans
; US-10-369-493-17249

Query Match      30.7%; Score 735; DB 15; Length 482;
Best Local Similarity 35.0%; Pred. No. 3e-50;
Matches 164; Conservative 90; Mismatches 209; Indels 6; Gaps 3;

QY 9 VNGEWSVNOIEILIPIDSSLGFPVPMTRREVDHAKAGREALPAMAAITVYERAOYL 68
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 12 ISDQWGHGQGSIFVIDPQTNHVIATVPATTEADAKAIEAKGCAIARQMPIHEBAAL 71
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 69 HKAADIERDKEEIAITVLAKEISKAYNASVTEVTRADLIYYAABEGIRLSTSADEGGM 128
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 72 KRTAEIVTRDRELFAQTIAMEGSKTIREARSEVRRCITETLLSGEBAKRLHGETTIPFSQM 131
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 129 DASTGH--KLAIVIRROPVGIVALAIPYNYPNVLSGSKIAPALIGNVMPKPTQGSVSG 186
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 132 ---PGERHRRVGYFRPPIGIVAITPFPNDPLNVAHKIGPALAAGNSLTIKSSFTPLLA 188
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 187 LVLAKEAFAGAGLPAGVFNITGRGSEIGDYIVHEEENFNFTGSPVQORIGKLAMGP 246
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 189 LRIYVLAKEAGLPKIVQVITGHSVITGPIITGHRVRLISFTGQYETGKIRSGGVKK 248
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 247 IMLELGKAGIVLADADLDNAKQIVAGAYDYSQGRCTAIKRYLVVEEVADELAEKISB 306
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 249 LAMEIGNSPTIYLQDAELMEAVAASCVSGAFCAAGNCIGVQRIYEGSVFNSFISFVA 308
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 307 NNAKLSVGPFPDNAT--VTPVIDNSADFIRESLVVDARQKAKELNFKDGLITGLTD 365
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 309 QTKQRLGPGKQSEETDIPMISSEKAKRIERWVEAKERBARVLTGRRGTGAYFEPTVLT 368
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 366 HTYLDKLAWEFPFGPILPPIRYKDAEBAVALANKSDPGLQSVFTRDOKAFDIANKLE 425
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 369 NVSPSSRLAKEERFAIVILEGVHSLTEAIARANDVDVDFGLQAGLTNNMLTAAPSALEKLE 428
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 426 VGTVAHNNKTGRGPDNFPFLGLKSGAGVQGIKRYSIAMTNNKSIIVLDM 474
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 429 VGGIWMNDSSDVVIDAMFPGGIIKKSGLGREGVYVAIEEMTEQKVVAFHL 477
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 32
; US-10-369-493-16454
; Sequence 16454, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
```

```

; SEQ ID NO 16454
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-10-369-493-16454

Query Match      30.4%; Score 728; DB 15; Length 470;
Best Local Similarity 35.7%; Pred. No. 1.1e-49;
Matches 167; Conservative 97; Mismatches 196; Indels 8; Gaps 5;

QY 8 VNGEWSVNOIEILIPIDSSLGFPVPMTRREVDHAKAGREALPAMAAITVYERAOYL 67
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 6 YINGDMKSVNTKPLAPVSEETLAEINAGTEDEVDKENTAKAMTKNTKNTLSADRAIT 65
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 68 LKAAADIERDKEEIAITVLAKEISKAYNASVTEVTRADLIYYAABEGIRLSTSADEGGM 127
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 66 LEKVAKQMDEREEREEFAEIIAKEAARPIRARGVDRVQTGYFAAEAKRI--YGETLP 122
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 128 MASTG--HKAIVIRROPVGIVALAIPYNYPNVLSGSKIAPALIGNVMPKPTQGSVS 185
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 123 LPAAPRADGRIAVYTRKPIGVIGALTFFNPLNVAHKGPALAGNTVVLAPADQTPIS 182
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 186 GLVLAKEAFAGAGLPAGVFNITGRGSEIGDYIVHEEENFNFTGSPVQORIGKLAMGP 245
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 183 SYALVBLFEAGLPNGALNIIISGPGSTVGEALVKNDVYASITFTGSPKVGIGIKQKAGLK 242
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 246 PIMLELGKAGIVLADADLDNAKQIVAGAYDYSQGRCTAIKRYLVVEEVADELAEKI 304
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 243 RVTLEIGSVAAYIIDEVDVLTDEIERYVMGAFVANNQVCISVQVVFHETKHEFLSKL 302
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 305 SENNAKLSVGP--PDNATVTPVIDNSADFIRESLVVDARQKAKEL--NEFKRDGLITGLT 362
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 303 KKAWSVAVGDDPLEETDVSALISRDVERIDMWQEALEKAGATVLCGKKRDARLEET 362
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 363 LFDHTYLDKLAWEFPFGPILPPIRYKDAEBAVALANKSDPGLQSVFTRDOKAFDIAN 422
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 363 VLAKNVPHKSVCCQVYFGLMTVNTFKEDVLEIQVNNSRYGLQAGVFNMLPKMARALD 422
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 423 KLEVTVAHNNKTGRGPDNFPFLGLKSGAGVQGIKRYSIAMTNNKSI 470
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 423 ELEVGVMINDIPTFRVDMFPGVYSGTGEGIKYVAIEEMTEMLV 470
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 33
; US-09-815-242-10264
; Sequence 10264, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl U.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
```

```

: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-07-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FASTSQ for Windows Version 4.0
: SEQ ID NO 10264
: LENGTH: 482
: TYPE: PRY
: ORGANISM: Escherichia coli
US-03-815-242-10264

```

Query Match	30.3%;	Score 726;	DB 9;	Length 482;
Best Local Similarity	34.3%;	Pred. No. 1.6e-49;		
Matches 164;	Conservative 92;	Mismatches 206;	Indels 16;	Gaps 7

Oy	6	QNYNNGWKSXSV--OIEHLSPIIDBSLFEVPAWTRREEDHAMKRGALPAMALTYVE	63
Db	12	QALINGEMWLDANNGEAIDVTNPNMGDKGSVFKMGADETFRALIDANBALPAMALTRKE	71
Oy	64	RAQYIHAQADIIRDKBEIATVYLAKISQAINASTVEVTRADLRVAAEGIRLSTSD	123
Db	72	RATLRNWFNMMEHODDLRLMTTEOGKPLAEAKGEISYASFIEMFAEEGKRIFYDPTI	131
Oy	124	EGGKMDASTGHKLAVIRBOPVGVILAAIPYNPVNLSSGKIALPALIGNVYMFPRPOGS	183
Db	132	PGHOAD-----KALIVKOPIGTALITWNPBPAAITRKAGPALAACCTVILPAPSOCTR	186
Oy	184	VSGVLAKAPAEAGLPAGVFENTITGSGSEIBDIYIEHEEWFNEINFSTGTPRGORIGKLAG	243
Db	187	FSALALAEALIRAGVPGVGVNVTGSAGAVNGELNSPLVRLSTSTGSTEIGRQLMQCA	246
Oy	244	--MRPIMLELGGKADGIVLADADLDNAAKOIYAGADYVSGORCTAIRVILVVEVADELA	301
Db	247	KDIKKVSLTEIGNAPFIVPDADDLDAVVALGALASKFRNAGQTCVCANLEYVQDGVYDRFA	306
Oy	302	EKISENVAKLISVGDPPDN--ATVPVYIDDSADFISLYVDAROKGAXEL---NEFKDGR	357
Db	307	EKIQAQVASKIHGDGLDNGVTIIPGLIDEQAIVAKVEENHLDALKEKARVAVCGKAKHERGN	366
Oy	358	LLTPGGLFVHTVLLMDKLAMEBPFGRILPIIRVYDAEBAVALANKSPFGLQSSVFTTRDOKA	417
Db	367	PFQGTILVDVPAKAVSKSEETFGPLALPLRFPRDEADVIAQANDTEFGLAALFYARDLSTRV	426
Oy	418	FDIANKLEEVGTVIHNKKTG--RGPDPNPFPLGLSGAGVQOGRISYEAMTNYKSVILDM	474
Db	427	FRGEEALEYGVIVIN--TGLISNEVAPFGGIIASGIGREGSKYTGIEDLEIKYVICIGI	482

```

RESULT 34
US-10-369-493-23511
; Sequence 23511, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23511
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-23511

```

Query Match	30.3%;	Score 726;	DB 15;	Length 482;
Best Local Similarity	34.3%;	Pred. No. 1.6e-49;		

Matches	164:	Conservative	92:	Mismatches	206:	Indels	16:	Gaps	7
Oy	6	QNVYNGEWKSSVN--QIEIILSPIDDSSIGFVPAMTREETVEHDAMKAGREALPAMALITYE	63						
Db	12	QALINGEWDLDANNEEALIDVTNPANGDKIGSVPKMGADDETRAIIDAANLALPAMRLTAKE	71						
Oy	64	RAQVLHRAADIIEEDKHEIATVLAKEISKAVNASTVEVTRADILRYAAEEGIRLSTAD	123						
Db	72	RATILRWENFMMEHODDLARMLTLEQKPLAEAKGEISYASFTLWFAEEGKRIYDGT	131						
Oy	124	EGKMDASTGHLKLVIRIQPVGIVLATAFVNYPVNLSSGKIAPALIGNVVMFKPPTGS	183						
Db	132	PGHQAD-----KRLIVIKQPIGVTRALITPMNFPAMITRKAGPALAACCTMVLKPSQTP	186						
Oy	184	VSLVLAKAFPAEAGLPAGVENVTTTGRGSEIGDYIVEHEEVNFIFETGSTPVGORIGLAG	243						
Db	187	FSALALAEALIRAGVPAGVFVNVVTSAGAVGNELTSPVLKLSPTGSTIRQLMEQCA	246						
Oy	244	--MRPIMLELGGKAGIYVLADADLDNNAKOIVAAAGYDYGSGRCRAIKRVLYVEEYADBLA	301						
Db	247	KDIKKVSLELGGNAPFVIFEDDADLDKAVEGALASKFRNAGQCCANLHYQDDGYDFFA	306						
Oy	302	EKISEBNYAKLSVGPFPDN-ATVTVVIDDNSADPIESLVDVROKAGAKEL--NEFKRDR	357						
Db	307	EKLQOAVSKHITGGLDNGVTITGPIDEKAVAKYBEHADLBEKARVYCGKAHERGN	366						
Oy	358	LITPGLPDHVTLDMKLAMEBPFGPIILPIIRYKDAEBAVALANKSDPGLQSSVFTDFOKA	417						
Db	367	FFQGTILVDVPANKKVSKEETFFGLAPLFRFKDEADYIAQANDTEFGLAAYFVARDLSRV	426						
Oy	418	FDIANKLEVGIVHINNKTG-RGPNPFPPLGKSGGAGYQIGIRYISEAMTNVKSIVLDM	474						
Db	427	FRVGEALLEYGIVGIN-TGIISNEVAPPGIKASGLGREGSKYGIEDYLKIKMYCIGL	482						
RESULT 35									
US-10-282-122A-56650									
; Sequence 56650, Application US/10282122A									
; Publication No. US20040029129A1									
; GENERAL INFORMATION:									
; APPLICANT: Wang, Liangsu									
; APPLICANT: Zamudio, Carlos									
; APPLICANT: Malone, Cheryl									
; APPLICANT: Haselbeck, Robert									
; APPLICANT: Ohlsen, Kari									
; APPLICANT: Zvekind, Judith									
; APPLICANT: Wall, Daniel									
; APPLICANT: Trawick, John									
; APPLICANT: Carr, Grant									
; APPLICANT: Yamamoto, Robert									
; APPLICANT: Forsyth, R.									
; APPLICANT: Xu, H.									
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms									
; FILE REFERENCE: ELITRA.034A									
; CURRENT APPLICATION NUMBER: US/10/282,122A									
; CURRENT FILING DATE: 2003-02-20									
; PRIOR APPLICATION NUMBER: 60/191,078									
; PRIOR FILING DATE: 2000-03-21									
; PRIOR APPLICATION NUMBER: 60/206,848									
; PRIOR FILING DATE: 2000-05-23									

```

; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56650
; LENGTH: 482
; TYPE: PR
; ORGANISM: Escherichia coli
US-10-282-122A-56650

Query Match
Best Local Similarity 34.3%; Score 726; DB 15; Length 482;
Matches 164; Conservative 92; Mismatches 206; Indels 16; Gaps 7;

QY 6 QNVNGEKSSVNV--QIEILSPIDSSLSGFVPAMTREEDVHAMKAGREALPMAALTVE 63
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 12 QALINGEWLDANGEMADVTNPANGDKGSVPMGADFRRAIDANRLPAMRALTAKE 71
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 64 RAQYLHKAADIERDKEEIAVLAKEISKAYNASVTEVVRTADLRVAAEGIRLSTSD 123
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 72 RATILRNFMNLMEHODDLARLMTLEQKPLAEAKGEISYASFIEWFAEGKRIYDIT 131
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 124 EGGKXDASTGHTLAVIRPQVGVLAIAFYNYPNVLSGSKIAPALIGNVWFKPTQGS 183
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 132 PEGHAD----KRLVYIKQPIGVTAITPMNFPMAITRKAGPALAAGCTWMLKPAQTP 186
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 184 VSGLVLAFAFAEGLPAGVFNTITGSGSEIGDIYVHEEWNFINFPVSGORIGKLAG 243
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 187 FSNALAEALAIRAGVPAAGVFNVTGSAGAVNELTNPVLRKLSFGSTGSEIRQLMEQA 246
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 244 --MRPIMLELGGKADGIVLADLDNAAKQIVAGAYDSGORCTAIKRVLVVEVADELA 301
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 247 KQIKKYSLELGGNAPFIVPDADLDKAVEGALASKFRNAGQTCVCANRLYVDDGYDRPA 306
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 302 EKISENVALSYGDPEDN--ATVTPVIDNSADPIESLVVDAROKAKEL---NEPKRDR 357
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 307 EKLQQAIVSKLHTGDGIDNVTIGPLIDERAIVAKVEHIALDEKGRVVGGAHERGN 366
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 358 LITPGLFHDVTLDMKLAMEEPFGPIPIIRVDAEBAVALANKSDPLGSSVFTDFOKA 417
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 367 FFGPTLVDPANAKVSKKETFGPLAFRFPDEADVIAQANDTEGLAAYRYADLSV 426
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 418 FDIANKLEVGVTHINNKTG--RGPDPNPFILGLKSGAGVQIRYSIEAMTNVKSIVLDM 474
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 427 FVVGELLEYGIVGIN--TGIIISNEVAPFGGIRKASGLGREGSKYGIEDYLEIKYMCIGL 482
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

RESULT 36
US-10-282-122A-46932
; Sequence 46932, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Fgreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848

```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46932
; LENGTH: 483
; TYPE: PR
; ORGANISM: Bacillus anthracis
US-10-282-122A-46932

Query Match
Best Local Similarity 35.7%; Score 721.5; DB 15; Length 483;
Matches 171; Conservative 82; Mismatches 203; Indels 23; Gaps 8;

QY 8 YNGWKSSVNDIELSPIDSSLSGFVPAMTREEDVHAMKAGREALPMAALTVERAQ 67
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 16 YINGEMITLQEQIEVNNPPTKEIFATVPKGVTEAKQVADAHAHEFKSKLTAADRAX 75
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 68 LHKADIERDKEEIAVLAKEISKAYNASVTEVVRTADLRVAAEGIRLSTADEGK 127
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 76 LKQWFTLIDENKEEIALMTKEQKFPALAEVAVYANSPWVYAEGRV-----YGE 130
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 128 MDASTGHTLAVIRPQVGVLAIAFYNYPNVLSGSKIAPALIGNVWFKPTQGSVGL 187
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 131 IPASHPNRIILVMKQPGVMAITPMNFPMAITRKVAPALAGCTAVVAPASQTPRLTXL 190
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 188 VLAKEFAERGLPAGVFNTITGSGSEIGDIYVHEEWNFINFPVSGORIGKLAG--MR 245
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 191 KLAELAHENDIPKGINIVTGSAKAIDTWMEDGVRKVSFGSTGSEIGELMASAQTK 250
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 246 PIMLELGGKADGIVLADLDNAAKQIVAGAYDSGORCTAIKRVLVVEVADELA 305
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 251 KVSLELGGHAPFIVMDDLDKAVEAVIGSKFRNAGQTCITNRVFPQDEVEAEVKE 310
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 306 ENVAKLSVGDPF--DNATYTPVIDNSADPIESLVVDAROKAKELNEFR---DGRLLT 360
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 311 KAVGQKTKGDKGDDGTTGPLIDENAVSKVQEHIEDAIQKGGTVLYGGQKVAELDGHFWQ 370
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 361 P---GLFPHVTLDMLKLANEPPFGPIPIIRVDAEBAVALANKSPFGLSSVFTDFOKA 417
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 371 FTVIGLADTMLCNM---EETFGPAPVAKFKETVEVERANRTHYGLAAYIFTDIOISQ 427
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 418 FDIANKLEVGVTHINNKTG--RGPDPNPFILGLKSGAGVQIRYSIEAMTNVKSIVLDM 474
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 428 FOISEALEYIGLNDGL--PSVAQAPFGGKESGIGREGGHFGLIEYLEIKYISLGL 483
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

RESULT 37
US-10-282-122A-60838
; Sequence 60838, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari

```

APPLICANT: Zysekina, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 60838
LENGTH: 488
TYPE: PRT
ORGANISM: Listeria monocytogenes
US-10-282-122A-60838

Query Match 30.0%; Score 718; DB 15; Length 488;
Best Local Similarity 33.8%; Pred. No. 7, 2e-49;
Matches 161; Conservative 96; Mismatches 201; Indels 18; Gaps 6;
QY 8 YVNGEMKSSVNO--LEILSPIDSSLGFPVPMTRREEDVHAKAGREALPAMAALTVERA 65
DB 16 PINGMCTOGDNKETDYNVPANGDVIYAKIAQGPSETKAIYKAADAFPMAMKMLADRV 75
QY 66 QYLHRAADIIRDEKEIATVLAKETISKAYNASVTEVVRTADLIRYAAEGIRLSTSADEG 125
DB 76 KLIHKTADIMEBKADTLAKIMTLEOQKPLKESKGEVLTGVENFRFAAEARL-----YG 130
QY 126 GKMDASTGHKLAIVIRQPGVIGLAPVNYVNLSSKIAAPLIGNVMPFPGQSGS 185
DB 131 EITFPNNNAF-IYKQPIGVAAITPMNFGGMVTRKLAAPALAGNTIVLPSGDTPLS 189
QY 186 GLVLAAPFAEAGLPAGVFNTITGRGSEIGDYIYHEBVENFINTGSTPGORIGKLA-- 243
DB 190 ALAIFETIEBAGLPBGVANIYVWSSKEIGETITLSDPVAKLFTGSTKYGQILFKQSADT 249
QY 244 MRPIMLELGGKAGIVLADADLDNAKQIVAGAYDYSQGRCTAIRKVLVEEVADELAEK 303
DB 250 LKKTISTELGHAFFIVFDANDDAVNDLVAAKFRNNGVGVSPRIFVAKETKEKFTKA 309
QY 304 ISENVAKLSVGPDPFNATYTPYIDNSADFIESLVVDAQKAKELNEKR-----DG 356
DB 310 IYAKVEQLKVNGLDGVNGLIRBDAIDKIDKQKNAEKGAKVLTGGRLTGSDYDKG 369
QY 357 RLITGLFDHVTLDKLAWEBPFGPIPLIRVKADEAVAIANKSDPFGQSSVFTTRDFOK 416
DB 370 NRYKFTVLDNVTARKDIFYEETFGVPIPLITEBTEBALEMANDSEFGLASFTYKDLAR 429
QY 417 AFDIANKLEVGTVHINNKTGRPDNFPFLGLKSGAGVQGIIRYSIEAMTNVXSIYL 472

Db 430 VEKVGALEYGWGANETIAISNPER-FPGGVKSHSGFGRENHGYGWEYIYQKFINL 484
RESULT 38
US-10-369-493-18199
Sequence 18199, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO: 18199
LENGTH: 493
TYPE: PRT
ORGANISM: Thermoplasma acidophilum
US-10-369-493-18199

Query Match 29.9%; Score 716.5; DB 15; Length 493;
Best Local Similarity 34.6%; Pred. No. 9, 6e-49;
Matches 167; Conservative 97; Mismatches 191; Indels 27; Gaps 9;
QY 8 YVNGEM--KSSVNOLEILSPIDSSLGFPVPMTRREEDVHAKAGREALPAMAALTVERA 65
DB 6 YIDGQWNNSSGKTVYDKSPYTGQVIGRFEATRDVDRAIDDAEDAWANNLDGVSERS 65
QY 66 QYLHRAADIIRDEKEIATVLAKETISKAYNASVTEVVRTADLIRYAAEGIRLSTSADEG 125
DB 66 KLIYRAKELIEKNRALENIIMEENGKPKVKAKEVDVIDQIYVAAEMARKLNGEVVEG 125
QY 126 GKMDASTGHKLAIVIRQPGVIGLAPVNYVNLSSKIAAPLIGNVMPFPGQSGS 183
DB 126 ----TSSHRKI FQYKYVYGIVALTPNFPAGVARTLAAPLITGNVTIVLPSSDTSGS 180
QY 184 VSGVLAAPFAEAGLPAGVFNTITGRGSEIGDYIYHEBVENFINTGSTPGORIGK--L 241
DB 181 AEWIV--RKFPVAGVPGKVLNFTGRGSEIGDYIYHEBVENFINTGSTPGORIGK--L 238
QY 242 AGMRPIMLELGGKAGIVLADADLDNAKQIVAGAYDYSQGRCTAIRKVLVEEVADELA 301
DB 239 ANMAKLTILELGGKAPFVWVXODMDNALKTLLMAKYVWNAQSCIAEBRLVYHEDIYTFM 298
QY 302 EKISENVAKLSVGPDPFNATYTPYIDNSADFIESLVVDAQKAK-----ELNEFK 353
DB 299 SRFVLSRKLLGSP-KNADMGPLINKGALQATSEIYVEABESGKILFGSQPSLGPY 357
QY 354 RDGRLITGLFDHVTLDKLAWEBPFGPIPLIRVKADEAVAIANKSDPFGQSSVFTTRD 413
DB 358 RNGVFFPLTIIGNADQSKITQEBIFAPVIGARKISSEVBYNDLANDSKYGLASLFLPKD 417
QY 414 FOKAEDIANKLEVGTVHINNKTGRPDNFP--FLGLKSGAGVQGIIRYSIEAMTNVXSIY 471
DB 418 PNIIFEASBRIRFGLLYVNP---GPEASQGVHTGFRMTGQAGBSKYGISEYLLKNIY 474
QY 472 LD 473
DB 475 VD 476
RESULT 39
US-10-369-493-18593
Sequence 18593, Application US/10369493
Publication No. US20030233675A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 18593
/ LENGTH: 480
/ TYPE: PRF
/ ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-18593

Query Match      29.8%; Score 713.5; DB 15; Length 480;
Best Local Similarity 35.0%; Pred. No. 1.6e-48;
Matches 165; Conservative 86; Mismatches 210; Indels 11; Gaps 6;

QY      8 YNNGEKSSVNOEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVERAGY 67
DB      13 FVAGEVWSGEPVPVSDLDGGTFASVAADTSDAERALSATGVAHDIRETTVPERVEM 72
QY      68 LHKADIIERDKKEIATVLAKESKAYNASVTEVVTADLIRYAEGRILSTSADEGK 127
DB      73 LSIADGIRREBELAEVIREAGKPISSARGEVQSAERPRVAGE-LRLHTEGRTG- 130
QY      128 WDASTGK--LAVIRQPVGIYLAIPYVYNLSSKTIAPALIGNVVFKPPTQGSVS 185
DB      131 --TTAGHEWQAIIVKHEPMGTVCITPVYVPLSTMALQVAPALAGNAVIVKPAKTPIS 188
QY      186 GIYLAFAFAAGIPAGVFTITGRGSEIGDIYVEHEVNFINTGSTPVGQRIKLAGR 245
DB      189 GAILADIADAGIPDAVNFVPESSVIDDPLASDARVAIAVMTGSSGAGEHVARQSGIT 248
QY      246 PMLTELKGDAGIYLAADLDNNAKQIVAGAYDSGORTCAIKRVLVVEVADLEAKIS 305
DB      249 RLHMEIGSAPAVIFEDDADDAADATAGSLKTAGQCSAVSRVLAHESYHDELVSRI 308
QY      306 ENVAKLSVGDPPD-NATVTPVIDNSADFIESLVVDAROKAKEINEFR---DG-RLIT 360
DB      309 DAAAEISIDLDFTDTTLGPLVSADQADWVAELVDADVGRATVVRGGRHVEDGVHYE 368
QY      361 PGLFDFHTLDMKLAWEPPGPIPIIRVYDABEAVAIANKSDGLOSVFTTDFOKAPFI 420
DB      369 PTLADVPRDARIIVDEEQFVCAVTVTDEDDAVRTANGSELALDAAVFTADHDAMRV 428
QY      421 ANKLEVGTVHINNKTGRGPDNFPFLGLKSGAGVQGIYRSIEAMTVKSIYL 472
DB      423 AARVANAAGVRIINGAPBSHGIDVPEFGNDASGIGREGIDSTIHFVREKSIILL 480

RESULT 40
US-10-282-122A-59544
/ Sequence 59544, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
```

```
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patencin version 3.1
/ SEQ ID NO 59544
/ LENGTH: 482
/ TYPE: PRF
/ ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59544

Query Match      29.7%; Score 711; DB 15; Length 482;
Best Local Similarity 34.1%; Pred. No. 2.6e-48;
Matches 163; Conservative 90; Mismatches 209; Indels 16; Gaps 7;

QY      6 QNYVNGEKSSVN--QIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVE 63
DB      12 QALINGQWDAENGVIAVTPNANGEOIGSVKMGADETREAIKANRALPAMRALTAKE 71
QY      64 RAOYHKAADIIRERKEEATVLAKESKAYNASVTEVVTADLIRYAEGRILSTSADE 123
DB      72 RANILIRFMDLMENQDILARLTLEQGPALAEAGEISYASFIEMFABEGRITYGDTI 131
QY      124 EGGKMDASTGHTLAVIRQPVGIYLAIPYVYNLSSKTIAPALIGNVVFKPPTQGS 183
DB      132 PCHQND-----KRLVITQPIGTVAITPWNFPAMITRKAGPALAGCTWVLPKASQRP 186
QY      184 VSGVLAKAFABAGIPAGVFTITGRGSEIGDIYVEHEVNFINTGSTPVGQRIKLAG 243
DB      187 FALMALAEIANAGIPAGVFNVTGSAGVGGELTNSPLVRLSTGSTEIGRLMEQCA 246
QY      244 --MRIMELEKGDAGIYLAADLDNNAKQIVAGAYDSGORTCAIKRVLVVEVADLE 301
DB      247 KDIKVSLELGGNAPVIFDDADLDKAVEGALASFRNAGQTCVCANRLYVDGYDRPA 306
QY      302 EKISENVAKLSVGDPF-DNATVTPVIDNSADFIESLVVDAROKAKEINEFR---RDGR 357
DB      307 EKLQAVETKRIQDGLQDQVTTGPIIDEKAVAKVEBHADALAKAKAVVTGKPHALGNN 366
QY      368 LITPGLFHTLDMKLAWEPPGPIPIIRVYDABEAVAIANKSDGLOSVFTTDFOKA 417
DB      367 FPGPTILVVVPSAKVAKKEETGGLPAPLFRFXDEADVIAQANDTEFGLAAYFARDLSRV 426
QY      418 PDIANKLEVGTVHINNKTG-RGPDNFPFLGLKSGAGVQGIYRSIEAMTVKSIYLD 474
DB      427 FRVGEALEYGIIGIN--TGIISTEVAPQGVASGLGRBGRKGIEDYIEIKYMCIGL 482

RESULT 41
US-10-369-493-18711
```

```

; Sequence 18711, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18711
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-18711

```

```

Query Match      29.7%; Score 710.5; DB 15; Length 489;
Best Local Similarity 35.9%; Pred. No. 2.9e-48;
Matches 167; Conservative 84; Mismatches 199; Indels 15; Gaps 6;

```

```

QY 8 YVNGEM--KSNVQCEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTVYERA 65
DB 15 YIDGIVPPRGDGTLEVDPPSTRELTEIPACTTDDVDAAYRAADACREMAADTPQERA 74
QY 66 OYLHKAADIIRDEKEEIAVLTAKEISKAYNASVTEVRTADLIRYAABEGIRLSTSADEG 125
DB 75 RIVAAASEILASHREVDVLLATESGTYTKGAASEFASVGITDEAASPTSMN-----G 129
QY 126 GKMDSATHKLAIVIRQPVGIVLAPVNYVNLGSKTAPALIGANNVMEFPPQGSYS 185
DB 130 DHKQSVBEKENIVREPPGVGIIISPMNPLNLSIRAVAPVAAGNSVILKPASSTPT 189
QY 186 -GLVLAFAFABAGLPAGVFNITGRGSEIGDYVEHEEVNFINFTGTPVGORIGLQAM 244
DB 190 GGLLAKVFEAGLVBGVNVTGKSGELGDRLAGHPADVSPFGSTAGIGQVAATAGE 249
QY 245 RPIW--LEIGKDAQIVLADADLDNAKQIVAGAYDYSQORCTAIKRVLVVEVADELA 302
DB 250 NLALPAMELGANNVHIVTEDADLDADVDSGVFSGFLHSGQICISINRHVHEDVDYVE 309
QY 303 KISENVAKLSVDPDPDNATVT--PVIDDNSADPIESLVNDAKQAKELNEFRDGLTLP 361
DB 310 RLTERAALPFTGTADADTVGPIIDESQRDQILDIYERTVDAGATLETGDHDLVVER 369
QY 362 GLFDHVTLDMLKLAWEPPFPIPIIRVKDAEBAVAIAKNSDPGLQSVFTRDFOKAFDIA 421
DB 370 TVLSATATMDMAACNEHRRGPAFVIRFSDDBAEVLAELANSTEGVLSGSAAGTGARADIA 429
QY 422 NKELEVGVHINKTGRGPDNPFPLGLKSGAGVQGIKRSIEAMTN 466
DB 430 DRIDGMIHINDQPVNQEPHLRFPGYKNSGLG---RYNGEYVLN 470

```

```

RESULT 42
US-10-369-493-13571
; Sequence 13571, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493

```

```

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13571
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Thermoplasma volcanium
US-10-369-493-13571

```

```

Query Match      29.5%; Score 705.5; DB 15; Length 473;
Best Local Similarity 34.6%; Pred. No. 6.9e-48;
Matches 167; Conservative 94; Mismatches 194; Indels 27; Gaps 9;

```

```

QY 8 YVNGEMKSSVA--QELSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTVYERA 65
DB 1 YIDGQWVDSNGDITTKNPFVTEGYIGFPAASRDVDVRAADAEADQAKMDLQSVRS 60
QY 66 OYLHKAADIIRDEKEEIAVLTAKEISKAYNASVTEVRTADLIRYAABEGIRLSTSADEG 125
DB 61 KIIVAKELIEENRKELENIILIQENGKPVKEAEVDGVLDMQVYAEWARKLNGEVVEG 120
QY 126 GKMDSATHKLAIVIRQPVGIVLAPVNYVNLGSKTAPALIGANNVMEFPPQGS 183
DB 121 ----TTHRRKIPOYKYVGTVALTPWNPFGAGVARKLAPALITGNTVILKPSDTPGS 175
QY 184 VSLGLAFAFABAGLPAGVFNITGRGSEIGDYVEHEEVNFINFTGTPVGORIGK--L 241
DB 176 AEWIV--KFIQAGVPKALNYITGRGSDIYIEHKVALLIMTGSTSGQRI MQAS 233
QY 242 AGMRPIMLELGKDAQIVLADADLDNAKQIVAGAYDYSQORCTAIKRVLVVEVADELA 301
DB 244 ANMAKMLELGKAPFVWVKDADINNALKTLLMAKYNWAGQSCTAABRLVYHEDYDTFI 293
QY 302 EKISNVAKLSVDPDPDNATVT--PVIDDNSADPIESLVNDAKQAK-----ELNFK 353
DB 294 NKPELSKRIIVGPP--SRADWGPLINTLAKTEVEEAESEGARILYKSGKDLSPY 352
QY 354 RDGRLTGLPDPHTLDMKLAWEPPFPIPIIRVKDAEBAVAIAKNSDPGLQSVFTRD 413
DB 353 KNGVFLPTIIDNDQSKIRQEBIFAVIARQVSDVEEYISLANDSKYGLASTLFTKD 412
QY 414 FQKADINKLEVGTVHINKTGRGPDNPF--FLGLKSGAGVQGIKRSIEAMTNVKSIV 471
DB 413 PNLIFEAERIRFGLYINMP---GPEASQYHNGFRLTGAGSGSRGILEYLIKNIY 469
QY 472 LD 473
DB 470 VD 471

```

```

RESULT 43
US-09-997-664-3
; Sequence 3, Application US/09997664
; Patent No. US20020151003A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Bassat, Arle
; APPLICANT: Catermole, Monica
; APPLICANT: Gatenby, Anthony A.
; APPLICANT: Gibson, Katherine J.
; APPLICANT: Ramos-Gonzalez, Isabel
; APPLICANT: Ramos, Juan
; APPLICANT: Sarislan, Sima
; TITLE OF INVENTION: Method for the Production of p-Hydroxybenzoate in Species of
; FILE REFERENCE: BC1018 US CIP
; CURRENT APPLICATION NUMBER: US/09/997,664
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/585,174
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3

```



```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 68209
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Pseudomonas putida
; US-10-282-122A-68209

Query Match      29.4%; Score 704; DB 15; Length 480;
Best Local Similarity 34.0%; Pred. No. 9,3e-48;
Matches 162; Conservative 91; Mismatches 209; Indels 14; Gaps 7;

QY 6 QNYVNGEMKSSVNV--QIELSPIDSSSGFVPMTRREVDHMKAGREALPMAALTYVE 63
DB 12 QYINGEWLADUNGOTIKVTNPATGEVIGTVPKMGTAETRRALTEADKALPMPRLTAKE 71
QY 64 RAOYLHKAADIERDKKEIATVLAKEISKAVNASYTEVVRTADLIRYAAEGRISTSD 123
DB 72 RSKAKLRMEELMIENODDLARLMTTEQKPLAEKEIYVASFIEWPAEAKRIYDIT 131
QY 124 ECGKMDASTGHLAVIRROQVIGVILAIADYNYVNLGSKLAPALIGNVVFKPPTQGS 183
DB 132 PEHQPD-----RLIYIKQPIGVTALITPMNPPAMITRKAGPALAAGCTVILKPSQTP 186
QY 164 VSGVLAKAFAPAGCPACVFNITITGRSGSIGDIYVEHEVNFINTGSPVQGRIG--L 241
DB 187 YSALLVELAHBRAGIPAGGLSVVTSAGVSGELTGNLSVRKLSFTGSGTEIGRLMEBCA 246
QY 242 AGMRPIMLELGGKDGIVLADADLNNAAKOIVAGAYDYSQGRCTAIKRYLVVEEVADELA 301
DB 247 KQIKVVSLELGGNAPPIVFDADLDKAVGAILIKTRNNGOTCVCANRIYVDGYIDARA 306
QY 302 EKISENVAKLSVGDPEFNATVT--PVIDNSADFIESLVVDAROKAKELNEFKR--DGRLL 359
DB 307 EKLLAAVAALAKIGNGIEBGTITGPIIDGKAVAKVQHIDAVSKAKVLGSGKLEGNF 366
QY 360 TPGLEDHTLDMKLAKEEFPGLPIIRKYDAEBAVALANKSDPGLQSSVFPDRDQKAD 419
DB 367 EFTLIVDPKTAIVAKETFGPLAPLFRFKDAEVLAMSNDEFGLASIFYARDMSRVFR 426
QY 420 IANKLEVGTVHINNKTGR--GPDNFPPLGKSGGAGVQGRISIEAMTVKSIVLDM 474
DB 427 VAEALIEGVNGIN--TGLISNEVAPFGIKASGLGRESKGTIEDLEIKYICISV 480

RESULT 46
US-10-369-493-17169
; Sequence 17169, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

```

```

; SEQ ID NO 17169
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Bacillus halodurans
; US-10-369-493-17169

Query Match      29.1%; Score 697.5; DB 15; Length 468;
Best Local Similarity 33.4%; Pred. No. 3e-47;
Matches 158; Conservative 93; Mismatches 209; Indels 13; Gaps 5;

QY 8 YNNGEMKSSVNOIELSPIDSSSGFVPMTRREVDHMKAGREALPMAALTYVERQY 67
DB 4 YNNGEMKSSVNOIELSPIDSSSGFVPMTRREVDHMKAGREALPMAALTYVERQY 63
QY 68 LHKADIERDKKEIATVLAKEISKAVNASYTEVVRTADLIRYAAEGRISTSDDEGCK 127
DB 64 LMRWQULIDBOODELGEIMTTEQKPLAEALGEVQYANSFIOWIAEAKRI-----YDIT 118
QY 128 MDASTGHLAVIRROQVIGVILAIADYNYVNLGSKLAPALIGNVVFKPPTQGSVGL 187
DB 119 IPASAINRRIIVQKQVGVIAAITPMNPPAMITRKAVPALAAGCTALVKDAEQPTPLAL 178
QY 188 VLAKAFAPAGCPACVFNITITGRSGSIGDIYVEHEVNFINTGSPVQGRIG--GLAAMR 245
DB 179 KLAQLAEEAGIPAGGLAVITGNADIGRAWLEDSRVRRKITPTGTEVGKLMRGAQTVK 238
QY 246 PIMLELGGKDGIVLADADLNNAAKOIVAGAYDYSQGRCTAIKRYLVVEEVADELAEKIS 305
DB 239 KISLELGGHAPFITMDNLEEAVDQVIAFRNAGQTCVCANRIYVAREIAEATEKFA 298
QY 306 ENVAKLSVGDPE--DNATVTPVIDNSADFIESLVVDAROKAKELNEFKRDGRLLTPGLF 364
DB 299 AKVVELKVGNGLEBGTITGPIIDKAAVEKAHIDALKKGGQTVVGGRWNTNHPFEPTII 358
QY 365 DHTYIDMKLAKEEFPGLPIIRKYDAEBAVALANKSPGLQSSVFPDRDQKAPIANKL 424
DB 359 TGAIDEMLCMNEETFGPLAPVATPDTEEEVERANHTPYGAAAYVTENIRALRLSEKL 418
QY 425 EYGVTHINNKTGRGPD--NFPPLGLKSGGAGVQGRISIEAMTVKSIVLDMK 475
DB 419 EYGVGVNDGM--PSVAQAPFGKSGSLGRBGGKGIIEYLEVYKYSVAFK 468

RESULT 47
US-10-369-493-18996
; Sequence 18996, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18996
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Anabaena PCC7120
; US-10-369-493-18996

Query Match      29.0%; Score 694; DB 15; Length 473;
Best Local Similarity 32.6%; Pred. No. 5,8e-47;
Matches 157; Conservative 102; Mismatches 198; Indels 24; Gaps 8;

QY 8 YNNGEMKSSVNOIELSPIDSSSGFVPMTRREVDHMKAGREALPMAALTYVER 64
DB 15 YNNGEMKSSVNOIELSPIDSSSGFVPMTRREVDHMKAGREALPMAALTYVER 64

```


Db 1 YINGCINATATELTINSHPADKEIVATFPRSQADDVDRVAAARQAVSGWRKVAPAR 60
Qy 65 AQLYHKAADIIRDRKEIATVLAKEISKAYNASVTEVATADLIRAAEGRLSTSADE 124
Db 61 AAYIFRVGELLQHKKEELAQLSIREKGPLETEARGVQGVCAFYSSAGEGRLLF----- 115
Qy 125 GSKMDASTGKLAVIRROPVGIYLAIAFVNYPVNLSGSKIAIPALIGANVVFEPFGQSV 184
Db 116 GQTTSEEMNKFMATRMKMGICVLCALITTPNFVAITCMKAPALVCGNTVILKPADDTA 175
Qy 185 SGVLAKAPAEAGLPAGVNTITGRSEIGDYIVEHEVNFINTGSTPEVQRIQLDQM 244
Db 176 CATKILIEIFAAAGLPFGVINLVHVGEEYKALVEHPNIDIVSFTSSATGAYVGETCR 235
Qy 245 --RPMLEIGKDGAGIVLADADLDNAKQIVAGANDYSGQRTAIXRVIVVEBVADELAE 302
Db 236 THKRCVLEWGGKNAQVWMDADLELDLGDALWGAFGTQCTATSRLLIHDIXEKFYT 295
Qy 303 KISENVAKLSVGDPPFNAT-VTPVIDNSADPIESIYVDAROKAKEL--NEPKDGR- 358
Db 296 MURETTSQRLDAGTEPEPTDIGITINNROLQVHEHYMTAREGAKILIGGEIATEGQLQ 355
Qy 359 ---LTPGLFDHYTLDMKLAWEPPGPIIPIRVKDAEBAVAIANKSDGLQSSVFTDP 414
Db 356 QGYFQPTLDVNTPGMVARBEIRFPVALIEVSTFEBAIALINDTKYGLSSVYTRDI 415
Qy 415 QKAFDIANKLEVGVTHINKTGRGDPNPFPLGKSGAGVQGIKRSIEAMTVKSIIVLDM 474
Db 416 NAFVAMRDIIEAGITVINGPTIGAEVHLDPFGVKQKGNHR-----EAGTALDVFTW 469
Qy 475 K 475
Db 470 K 470

RESULT 48
US-10-369-493-8741
Sequence 8741, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Hinkle, Gregory J.
APPLICANT: Cao, Yongwei
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 8741
LENGTH: 477
TYPE: PRT
ORGANISM: Ralstonia metallidurans
US-10-369-493-8741

Query Match 29.0%; Score 694; DB 15; Length 477;
Best Local Similarity 35.5%; Pred. No. 5,9e-47;
Matches 173; Conservative 83; Mismatches 197; Indels 34; Gaps 10;
Qy 5 YONYVNGEWSKSVNOIELSPIDDSIGFVPMTREEVN-HAMKAGREALPA----- 55
Db 3 YANYIDGECDSASG---RTVDNVN---PADTDIVSRHAASDRDAAAIVAAAAA 54
Qy 56 --WAALTTERAOYLHKADITERRDKEEATVLAKEISAVNASTVEVTRDLIRYAAE 113
Db 55 DQMKTPIGKRAKIINDAAHLEAANDTATETLREEGALMLARDEVVRSQOTRFYAV 114
Qy 114 EGRISTSDGEGKMDASTGHLAVIRROPVGIYLAIAFVNYPVNLSGSKIAIPALIGAV 173

Db 115 EGQTSGESGYNDPDD-----MLVYSLEPPLGVTVIISPWNFPVSIIPAKKIAPALITGNT 169
Qy 174 VMFKPPTQSGVGLVLAQAPAEAGLPAGVNTITGRSGEIGYIYHEHEVNFINTGSTP 233
Db 170 VVFKPSSAPLSGSLAEAFVAGIPIKGYLNLTSAAEAVGTIYVESRVAVSFTGSTS 229
Qy 223 VQGRIGKLAKMP-IMLEIGKDGAGIVLADADLDNAKQIVAGANDYSGQRTAIXRVIV 292
Db 230 ABEQIHKSVPMTTRQOMELGKRNPLIWMEDADLDRAVDLAVGGISLSQACTGTSRIIV 289
Qy 293 VEEVADLEAKISENVAKLSVGDPPD-NATVTPVIDNSADPIESIYVDAROK-----GA 346
Db 290 MBDVQAFTKLVAVVKALKIGSGMTAGMDLGPATRKQLETVLVYIAIGKQEAETLLCGG 349
Qy 347 KELAEPK-RDGRLLTPGLFDHYTLDMKLAWEPPGPIIPIRVKDAEBAVAIANKSDPGL 405
Db 350 EQLSDGDFAHGYVVAFTVFTDTQSMRIAREEIRFPVALIIEVDYADAIAQANDTEYGL 409
Qy 406 QSSVFTDRQKAFDIANKLEVGVTHINKTGRGDPNPFPLGKSGAGV--QGIRYSIEA 463
Db 410 SAAIATRRMRWMDFTRIDIESGTVKINRTTGNLVNAPPGLKRSSTTFRESGAGLEF 469
Qy 464 MTNFKSI 470
Db 470 YIQIKTV 476
RESULT 49
US-10-282-122A-69547
Sequence 69547, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 69547
LENGTH: 480


```

; FILE REFERENCE: ELITRA_034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75217
; LENGTH: 482
; TYPE: PRF
; ORGANISM: Salmonella typhimurium
US-10-282-122A-75217

```

```

Query Match      28.9%; Score 692; DB 15; Length 482;
Best Local Similarity 33.3%; Pred. No. 8.6e-47;
Matches 159; Conservative 95; Mismatches 208; Indels 16; Gaps 7;

```

```

QY 6 QNYNGEMKSSV--VNOIEILSPIDSSLGFPVPMTEBVDHMKAGREALPMAALTVE 63
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 12 QAFIDGAWVDADNGQTIKKNMPATGEIIGTVPRKGAETRAIEADKALPAMRALTAKE 71
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 64 RAQYLHKADIIERDDEIATVLAKEISKVNASVTEVVRTADLIRYAABEGIRLSTSD 123
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 72 RANILRWRFEIMTEHODDLARLMTLEQKPLAEAKGEIYAAASFIEWFAEBEKRIYGDVI 131
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 124 EGKMDASTGHLAVIRROPGVIGVLAIPYNYPVNLGSKIAIPALIGNVNMEKPTQGS 183
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 132 PGHQAD----KRLVITKQPIGVTAITPWNPFSAKITRAGPALAAGCTMVLKPAQTP 186
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 184 VSGVLVAKAFAPAGLPAGVFNITITGRGSEIGDIYIVHEEVENFTNGSTPVGQRIGK 243
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 187 FSAIALAEIAQRAQIPAGVFNVTGSGADIGSELTSNPLVRKLSFGSTIEIGRLMEQCA 246
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 244 --NRPIMLELGGDAGVILADADLNAKQIVAGAYDSGORCTAIKRVLYVEEVADELA 301
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 247 KQIKKVSLELGGNAPIVFDADLDKAVGALISKRRNAGQTCVCANRLVUDGVYDRFA 306
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 302 EKISENVAKLSVGDPPFD-NATVTPVIDNSADFIESLVVDARQKAKEL---NEFRDGR 357
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 307 EKNQAVNKLAVGDDGADVAIGPLIDEKAVAKVOEHIALDKGARVITGSAHLGCGN 366
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 358 LLTPGLFDHYTLDMKLAMEEPFGIILPIIRYKDAEENALANKSDPGLQSSVFTRPQKA 417
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 367 FPGPTLIADVPDNAKAKAETFEGLPLPLFRFSDEADVIRQANDTEFGLAAFYARDLSRV 426
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 418 FDIANKLVEGTVINNKTKG-RGPDNPFGLKSGGAGVOGIRYSIEAMTVKSIIVLDM 474
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 427 FRVGEALFYIGVIN--TGLISNEVAPFGSIKASGLGRBSKXGIEDYLEIKIMCIGL 482
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 52
US-10-369-493-13957
; Sequence 13957; Application US/10369493
; Publication No. US20030233675A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13957
; LENGTH: 478
; TYPE: PRF
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-13957

```

```

Query Match      28.9%; Score 691; DB 15; Length 478;
Best Local Similarity 33.5%; Pred. No. 1e-46;
Matches 159; Conservative 92; Mismatches 209; Indels 14; Gaps 7;

```

```

QY 6 QNYNGEMKSSV--QIEILSPIDSSLGFPVPMTRBVDHMKAGREALPMAALTVE 63
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 11 QAFIDGAWVDADNGQTIKKNMPATGEIIGTVPRKGAETRAIEADKALPAMRALTAKE 70
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 64 RAQYLHKADIIERDDEIATVLAKEISKVNASVTEVVRTADLIRYAABEGIRLSTSD 123
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 71 RAGKLRWFELMTIEHODDLARLMTLEQKPLAEAKGEIYAAASFIEWFAEBEKRIYGDVI 130
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 124 EGKMDASTGHLAVIRROPGVIGVLAIPYNYPVNLGSKIAIPALIGNVNMEKPTQGS 183
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 131 PGHQD----KRLVITKQPIGVTAITPWNPFSAKITRAGPALAAGCTMVLKPAQTP 185
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 184 VSGVLVAKAFAPAGLPAGVFNITITGRGSEIGDIYIVHEEVENFTNGSTPVGQRIGK-L 241
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 186 FSAFALAEIAQRAQIPKGFVSAGDIGSELTSNPLVRKLSFGSTIEIGRLMESECA 245
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 242 AGRPIMLELGGDAGVILADADLNAKQIVAGAYDSGORCTAIKRVLYVEEVADELA 301
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 246 KQIKKVSLELGGNAPIVFDADLDKAVGALISKRRNAGQTCVCANRLYUDGVYDRFA 305
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 302 EKISENVAKLSVGDPPFDNATVT-PVIDNSADFIESLVVDARQKAKELNEFK-RDGRLL 359
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 306 EKLKVAVALAKLIGNGLEBGTGTPGLIDEKAVAKVOEHIALDAVSKATVLAGKPMEGNPF 365
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 360 TGLFDHYTLDMKLAMEEPFGIILPIIRYKDAEENALANKSDPGLQSSVFTRPQKAD 419
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 366 EPTIILNPKDAVAVAKEETFGPLPLFRFXDADVIAMSNDEFGLASFYARDIGRVR 425
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 420 INKLEVGTVHNNKTKG-RGPDNPFGLKSGGAGVOGIRYSIEAMTVKSIIVL 472
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 426 VAEALFYGVNVA--TGLISNEVAPFGSIKASGLGRBSKXGIEDYLEIKIYCL 477
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 53
US-10-282-122A-51224
; Sequence 51224; Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.

```

```

/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See file wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 51224
/ LENGTH: 489
/ TYPE: PRT
/ ORGANISM: Bordetella pertussis
/ US-10-282-122A-51224

Query Match      28.8%; Score 690; DB 15; Length 489;
Best Local Similarity 33.4%; Pred. No. 1,3e-46;
Matches 161; Conservative 93; Mismatches 200; Indels 28; Gaps 8;

QY      8 YNNGEKSSVNOIEILPIDSSISG---FVPAMTSEVDHAKAGREALPAAALTYVE 63
      19 YIDGKRVGANGNST--PYDNPSTGKTIISVPLGSKKEAQAIDAAEALPAMSATGGE 76
DB      64 RAQYLHKAADIIERDKEEIAATYLAKEISKAYNASVTEVVRTADLIRYAAEGIRLSTSD 123
      77 RAAILLKMAQMMENQODLAVIMTSEQKVPPEAGEIYAASFLFMPAEAKRI----- 131
QY      124 EGGKMDASTGHTLAIVRRQPVGIVLAIAPYNYVNLSSGKIAPALIGNVVMFKPTQSS 183
DB      132 DDIIQSPKGGORLWMLKOPIGVTAITPWNFPAAMITRKVGPALAAGCTMVVKPQAQTP 191
QY      184 VEGLVIAKAPAEAGLPAGVFNITTGSGSEIGDYIVHEEYVNFNFTGSTPVGORIKLAG 243
DB      192 LRALALAVLAVLAGVAGVAFQVVTGSSREIGALCESEVVRKLSFTGSEIVGRTLMEOCA 251
QY      244 --MRPIMLELGGKADGIVLADADLDNAAKQIVAGAYDYGQRCCTAIKRYLVVEEVADELA 301
DB      252 PTKIKSLTELGNAPRIVFDADLDRAIDGILASKYRNAGQTCVCANRITYVQGVYEVA 311
QY      302 EKISEVVALISVGDPPD-NAITYPVVIDNSADPISLVDAROKAKEL---NEFRDGR 357
DB      312 KSLVEEVNMMKVGDDGEGVTOGPLLDTAAVAKVOEHADATAGAKVLAGEPHALGGS 371
QY      358 LITPGLFDVTLDMKLAWEPPGPIPIIRVKDAEAAVAIANKSDGLQSVFTPFOKA 417
DB      372 FPOPTVVRVTOGSMRAVEETGCPVAPLTKFSEDEVIGMANNTITGLAAYFYTRROYARI 431
QY      418 FDIANKLEVGTVHIN-----NKTGRGPDNFPFLGLKSGAGVQGIIRYSIEAMTNVKSIVL 472
DB      432 WRVSEALEYGIIVGINIGIISNEVG-----PFGVGVSQGLGRGSGSYGIEDYELKYLVCV 485
QY      473 DM 474
DB      486 DL 487

```

```

RESULT 54
US-10-369-493-11184
/ Sequence 11184, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xiandeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 11184
/ LENGTH: 480
/ TYPE: PRT
/ ORGANISM: Methanosarcina mazei
/ US-10-369-493-11184

Query Match      28.7%; Score 687; DB 15; Length 480;
Best Local Similarity 32.1%; Pred. No. 2.2e-46;
Matches 155; Conservative 105; Mismatches 207; Indels 16; Gaps 6;

QY      1 LTKXYONYNGEWM--SSVNOIEILPIDSSISGFPAMTSEVDHAKAGREALPAA 58
      3 MVQBYKLFGEFQKOSTGETFEDINPATLNLATIQVAGADVDMAVEAAEAGFRLWNE 62
DB      59 LTVYRAQYLHKAADIIERDKEEIAATYLAKEISKAYNASVTEVVRTADLIRYAAEGIRL 118
      63 VPAARRAVLFPAAAIILQERKEEISVMTKEMGKVLPTRGDVQDAIDITYAAEGGRM 122
QY      119 STSADEGKMDASTGHTLAIVRRQPVGIVLAIAPYNYVNLSSGKIAPALIGNVVMFKP 178
DB      123 L-----GETTSELKEKCMILIRPIGVGLITPNNFPIALPAMKVMALVAGNAIVFKP 177
QY      179 PTQGSVGLVIAKAPAEAGLPAGVFNITTGSGSEIGDYIVHEEYVNFNFTGSTPVGORI 238
DB      178 ASDTLAFKLIEVSEAGLPPGVINLVYTGCGTVGKAVVOHPRIKALSFTGSLDTGKWI 237
QY      239 GGLAG--MRPIMLELGGKADGIVLADADLDNAAKQIVAGAYDYGQRCCTAIKRYLVVEV 296
DB      238 MEECSKTKMRSVLELGNKVPVIYMDADLELALBGVLMGARFQTTGQRCCTATSRLILHEKI 297
QY      297 ADELAEKISENVAKLSVGD- FDNATVTPVIDNSADPISLVDAROKAKELNEFKR- 354
DB      298 KDEFTKRLLAARAKALSTIGGLLPETIDIPVINKQLEIKERYVKIKREKATILYGNRI 357
QY      355 ----DGRLLITPGLFDVTLDMKLAWEPPGPIPIIRVKDAEAAVAIANKSDGLQSSVF 410
DB      358 DPGLVGYFEBPFIIFDVAPDMRIAGEEIFGPVLGIFTSDLEAATILANSTKYGSIAIY 417
QY      411 TRDPQKADPIANKLEVEGVVHINKTGRGPDNFPFLGLKSGAGV- QGIIRYSIEAMTNVKS 469
DB      418 TGNIGNAARAIEKVBAGITVVAPTIGAIVHLFPGVGVTGNGFREGAETEAIVEKPSYKVA 477
QY      470 IYL 472
DB      478 VTI 480

```

```

RESULT 55
US-10-369-493-8312
/ Sequence 8312, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei

```


Db 1 GERVTRPRLVDVDPYTGTRVGTVPAS--VDDVRAAEVAAAYOALTRYERSQILER 57
Qy 71 AADIERDEKEIATVLAKEISKAYNASVTEVVRTADLIRYAAEGIRLSTADEGKMDA 130
Db 58 AASIMRERTEBAASDLISLSEGLSKODSRYEIGGVADVLRKFASEALR-----DDGSPFC 112
Qy 131 ST--GHKLAIVT-RROPV-GIVLAIAPIYVYVUNLSSGKIAPALIGNVVFKPPTQGSV 185
Db 113 DLTPHGKKRRVFSQREPRLAGIVAIITPFVHPMNQVAKIAPALATNNRVLTKPSEKVPUS 172
Qy 186 GLVLAFAFAAGIPAGVFNITGRGSEIGDIYVEHEVNFINTGSTPVGQRIKLAGNR 245
Db 173 AYLVDLVLEAGIPAPMLQVLTGDPREDLITPLAELVFTTGVAIGKTIARAAYR 232
Qy 246 PIMLEIGKDGAGIVLADADLDNNAKOIVAGAYDSGORCTAIRVILVVEEVADELAEKIS 305
Db 233 RYVLEIGGNDPLIVDDADLEERATLAOGSYNGSORCTAVRMLVQGSVAADFDLVV 292
Qy 306 ENVAKISVGDPPD-NATVTPVIDNSADPIESLIVDAROKAKEINEFKRDGLLTPGLF 364
Db 293 EKTRATFGDPDPAQSOMGTVIDVAALQFEARVNEAVASGARLLTGNQNGALVAPTVL 352
Qy 365 DHVTLDMKLAMEPRPILPIIRVKAEEVAIAANSDDGLOSVTFRPQKAFDIANTL 424
Db 353 DGVDPSTVLVREBTFGPVSPILTFDTLDALIRISNGTAFGLSSGVCVTRNODAITREFINEL 412
Qy 425 EVGTVAHNNKTRGSPNPFPLGLKSGAGV-OGIRYSIEAMTNVKS 469
Db 413 RVGTVAWEVPGRIELTFPGIKDGLGKEGVQAMKSFTNLKT 458

RESULT 58
US-10-369-493-17279
; Sequence 17279, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17279
; LENGTH: 485
; TYPE: PRN
; ORGANISM: Bacillus halodurans
US-10-369-493-17279

Query Match 28.5%; Score 682.5; DB 15; Length 485;
Best Local Similarity 33.4%; Pred. No. 5.1e-46;
Matches 159; Conservative 96; Mismatches 210; Indels 11; Gaps 5;
Qy 6 QNYVNGEW--KSSVNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVE 63
Db 7 RQFLAGEWIEGASTQITTYNNPYTGKEIHEIKASIDDDQAVRAKNAQKEWEGTLPQAQ 66
Qy 64 RAQVYHKAADIIRDEKEIATVLAKEISKAYNASVTEVVRTADLIRYAAEGIRLSTAD 123
Db 67 KQVLEKKAALALAEKQDEVIETLVNAGSSITIKASIEWGATLQITVVAALTFPLRM---- 121
Qy 124 EGGKMDASTGHKLAIVIRQPVGIVLAIAPIYVYVUNLSSGKIAPALIGNVVFKPPTQGS 183
Db 122 EGGILPSPNIPGKRNIRYSAKGVIGVISPFPNPLVLAAMSVAPALATGNAVVLKSCSDAP 181
Qy 184 V-SGLVLAQAFAPAGIPAGVFNITGRGSEIGDIYVEHEVNFINTGSTPVGQRIKLA 242

Db 182 ITSGLLIAELFEAEGLPRGVINNVVKAKEIGDAFVTHPIPKLISFTGSTEVGRHIAOLA 241
Qy 243 G--MRPIMLEIGKDGAGIVLADADLDNNAKOIVAGAYDSGORCTAIRVILVVEEVADEL 300
Db 242 ARELKEVLEIGGNNVMLVDDADLEKAAEKAAYKFLHQOICMALNRIIVDASITYSF 301
Qy 301 AEKISENVAKLSVGDPPDNAT-VTPVIDNSADPIESLIVDAROKAKEINEFKRDGLL 359
Db 302 VEVFKKVSQLOQTPAPAPATLIGPLINYKOIGRIQLVKESVAQGAVKVLEGHVQNLIM 361
Qy 360 TPGLFDHTLDMKLAMEPRPILPIIRVKAEEVAIAANSDDGLOSVTFRPQKAFD 419
Db 362 SPTLISEVTNMPVAKKEIPGPIAIIKAKQEAERAIALANDSPYGLSGSIFGSIHRVQ 421
Qy 420 IANKLEVGTVAHNNKTRGSPNPFPLGLKSGAGVQGVQIRYSIEAMTNVKSIVLDMK 475
Db 422 VAKQIDTMHINVDQPVNVEAHISIGGKDSIGRFGGEWVLDKTTYKWSIQEK 477

RESULT 59
US-09-815-242-14047
; Sequence 14047, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Twack, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14047
; LENGTH: 482
; TYPE: PRN
; ORGANISM: Salmonella typhi
US-09-815-242-14047

Query Match 28.5%; Score 682; DB 9; Length 482;
Best Local Similarity 33.1%; Pred. No. 5.5e-46;
Matches 158; Conservative 95; Mismatches 209; Indels 16; Gaps 7;
Qy 6 QNYVNGEKSS--VNOIEILSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTYVE 63
Db 12 QAFIDWRDADAGGVIVSPNANKP/GVNPQKGAESTRDAINANRALPAMRALTAKE 71
Qy 64 RAQVYHKAADIIRDEKEIATVLAKEISKAYNASVTEVVRTADLIRYAAEGIRLSTAD 123
Db 72 RANILRRFNNLMHQQDDLARMTLBOGKPLAAKGEISYASFTLEWFAEBSKRIYGTI 131
Qy 124 EGGKMDASTGHKLAIVIRQPVGIVLAIAPIYVYVUNLSSGKIAPALIGNVVFKPPTQGS 183

```

Db      132 PGHQT-----KRLVTKQPIGVTAATPWNPPSAMITRKAGPALAAGCTMVLKPSQRP 186
Qy      184 VSGLVAKAFAGELPAGVNTTGRSGEIGDIYVEHEVNFNPGSTPVGCRIGKLAG 243
Db      187 FSLALAEALQRAIGLAGVNVVTSAGDIGELTSPVLVRKLSFTGSTEIGQLMEQCA 246
Qy      244 --NRPIMLEGKDGAGIVLADADLNAKOIVAGAVDSGCRCTAIKRVLYVEEVADELA 301
Db      247 KDKKVSLELGNAPFIVFPDADLDKAVEGALASKFRNAGQTCVCANRLYVDGVDRFA 306
Qy      302 EKISENVAKLSVGDPPD-NATVTPVIDNSADFIESTLVVAROKAKEL---NEFRDGR 357
Db      307 EKLNQAVNKLAVDGLQADVAIGPLIDEKAVAKVQEHADALEKGAARVITGGBAHLKGN 366
Qy      358 LITPGELFDHYTLDMKLAMEEPFGPIPIIRVKAEEAVAIANKSDPGLQSVTFTRDFOKA 417
Db      367 FPGPTLIADVPDNAAKAKETFGPLAPLFRPSDEADVIRQANDTEFGLAAYFYARDLSRV 426
Qy      418 FDIANKLEVGTVHINNTKG-RGPDNPFPLGLKSGGAGVOGIRYSIEAMTVKSIYLDM 474
Db      427 FRVGEALFYGVGIN-TGIISNEVAPFGGIKASGLGREGSKYGIEDYLEIKYMCIGL 482

```

RESULT 60

```

US-10-282-122A-75920
/ Sequence 75920, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haseibeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Foreyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 75920
/ LENGTH: 482
/ TYPE: PRT
/ ORGANISM: Salmonella typhi
US-10-282-122A-75920

```

```

Query Match      28.5%; Score 682; DB 15; Length 482;
Best Local Similarity 33.1%; Pred. No. 5,5e-46;
Matches 158; Conservative 95; Mismatches 209; Indels 16; Gaps 7;

Qy      6 QNVNGEMKSS--VNOIEILSPIDSSLGFPVAMTREEDVHAMKAGREALPMAALTYVE 63
Db      12 QAFIDGDMADAGGVITPVSNPANGKPLGNVFMKGEETFRDAIINANRALPMAALTAKE 71
Qy      64 PAQYHKAADIIBERDEEATVLAKEISKANYASTVEVTRIDLIRYAAEEGRILSTSD 123
Db      72 RANITFRMNMIMHEODDLARLMTLEQGRPLAEAKGEISYASFIEMFAEBGRRIYDGI 131
Qy      124 EGGKMDASGHKLAIVRRQPVGVILAIPYNPVNLGSKTIAPALIGVWMEKPTQGS 183
Db      132 PGHQT-----KRLVTKQPIGVTAATPWNPPSAMITRKAGPALAAGCTMVLKPSQRP 186
Qy      184 VSGLVAKAFAGELPAGVNTTGRSGEIGDIYVEHEVNFNPGSTPVGCRIGKLAG 243
Db      187 FSLALAEALQRAIGLAGVNVVTSAGDIGELTSPVLVRKLSFTGSTEIGQLMEQCA 246
Qy      244 --NRPIMLEGKDGAGIVLADADLNAKOIVAGAVDSGCRCTAIKRVLYVEEVADELA 301
Db      247 KDKKVSLELGNAPFIVFPDADLDKAVEGALASKFRNAGQTCVCANRLYVDGVDRFA 306
Qy      302 EKISENVAKLSVGDPPD-NATVTPVIDNSADFIESTLVVAROKAKEL---NEFRDGR 357
Db      307 EKLNQAVNKLAVDGLQADVAIGPLIDEKAVAKVQEHADALEKGAARVITGGBAHLKGN 366
Qy      358 LITPGELFDHYTLDMKLAMEEPFGPIPIIRVKAEEAVAIANKSDPGLQSVTFTRDFOKA 417
Db      367 FPGPTLIADVPDNAAKAKETFGPLAPLFRPSDEADVIRQANDTEFGLAAYFYARDLSRV 426
Qy      418 FDIANKLEVGTVHINNTKG-RGPDNPFPLGLKSGGAGVOGIRYSIEAMTVKSIYLDM 474
Db      427 FRVGEALFYGVGIN-TGIISNEVAPFGGIKASGLGREGSKYGIEDYLEIKYMCIGL 482

```

RESULT 61

```

US-10-369-493-16681
/ Sequence 16681, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 16681
/ LENGTH: 460
/ TYPE: PRT
/ ORGANISM: Bacillus thuringiensis
US-10-369-493-16681

```

```

Query Match      28.4%; Score 680.5; DB 15; Length 460;
Best Local Similarity 34.6%; Pred. No. 6.8e-46;
Matches 161; Conservative 82; Mismatches 205; Indels 17; Gaps 6;

Qy      19 QTEILSPIDSSLGFPVAMTREEDVHAMKAGREALPMAALTYVEPAQYTLHRAADIIBERD 78
Db      4 QIEVNNPATKEIPATYFKGCVTEBAKQAVDAABAFKTMKSLTAADRAATLTKKFMPLIDGN 63
Qy      79 KEIATVLAKEISKANYASTVEVTRADLIRYAAEEGRILSTSDGGRKMDASTGHKLAIV 138
Db      64 KEIATVLAKEISKANYASTVEVTRADLIRYAAEEGRILSTSDGGRKMDASTGHKLAIV 118

```

[illegible]

```

RESULT 63
US-10-424-599-162810
/ Sequence 162810, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 162810
/ LENGTH: 537
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MFT3847_118035C.1.pep
US-10-424-599-162810

```

Query Match	28.3%	Score 677	DB 15	Length 537
Beet Local Similarity	34.0%	Pred. No. 1.6e+05		
Matches 161	Conservative	90	MissMatches 248	Indels 14
				Gaps 6
QY	6	ONYVNGEKKSVN--QIEILSPIDSSIGFVAMTRREEDHAKMGREALPAMALTYE	63	
DB	63	OGILGGKSDAIDGKTIKYNPATGESIVDAOCMGRETNDIAISAAYDYGSSKTTIAAE	122	
QY	64	PAQYLHKADIIERDKEEIATLAKESIKVAAVSYTEVVRTADLIRYAABEGIRLSTSD	123	
DB	123	RSKILRKMYDILMHKEELAQILTEQCKPLKESGEIYVYAGGLEFAAEKKRI-----	177	
QY	124	BGGKMDASTGHLKAVIRROPVGIYVLAIAFNYPVNLSSGSKIAPALIGANVVMFKPPTGS	183	
DB	178	YGDIVAPAPSPDRFLFVLKOPVGWVGAIIIPWNPFLMIRKVGAPALACCTVVIKPSLTP	237	
QY	184	VSGVLAKAFAMAGLPACVGFENTITRGSEIEGIIYIEHEEVNFINTGSTPVGQRI--GKL	241	
DB	238	LTALAABELPIQAGLIPGVVNVVMNAPDIGALLASDPVRKITPTGSAVAGKMLMAESA	297	
QY	242	AGMRPIMELTEGGKADGIVLADADILMAAKQIYAAGYDVSQGCSTAIKVLVYEEVADELA	301	
DB	238	EYVKKVSLELGGNACIYFPDDADLVAAKGTILAAKFRNSGQTCVCANRIYVOEGIYEKFA	357	
QY	302	EKISENVAKLVSQDEFP-DNATVTPVIIDNSADFISLIVDARQKAKEINTEFKRGRILLT	360	

Db 358 NALRDYANQNNKKGDDGFSEBVSQGPLINEAAVKVESLTHDTSKGAUYILJGKRHSIGLT 41 7.7
Qy 361 ---PGLFDHTLLDMKLAAMEBPFGLPLPIIRVADAEVAIAINKSDFGHOSSEVPTRDPOKA 41 7.7
Db 418 FYEPTYISDVNSPMHISREAFGVAAPLPLRFTEBEAIRINDNPAGIGSVVFNLSIORS 47 7.7
Qy 418 FDIANKLEVTGTHIINKTKRGPDNPFGLKSGAGVUGIYISIEAMTNVKS I 47 0
Db 478 WRVAELAEVLGVNVEGV-ISTEVAAPFGGFGKSGIGRGSGSTYCGMDEYILIKYV 52 9

```

RESULT 63
US-10-369-493-4453
/ Sequence 4453, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfang
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 4453
/ LENGTH: 491
/ TYPE: PRT
/ ORGANISM: Burkholderia fungorum
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(491)
/ OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-4453

```

Query Match	Similarity	28.3%	Score 676.5	DB 15	Length 491
Best Local	Similarity	34.1%	Fred. No. 1.6e-45		
Matches	Conservative	85	Mismatches 203	Indels 25	Gaps 10
Db	9	VNGEKKSVNO-IELISPIDDSSLGFWANMREVEYDHAKGREAL--PAMAAITVERA	65		
Db	18	VGDERSNGSGSFESINPIDGSVAITTAQASBSVDVDAVRAARQLETPARMDKVNHERG	77		
Qy	66	QYLRKADIIRDKKEIATVLAKETSKYNASVTEVTFADLIRYABEGIRLSTSBDBG	125		
Db	78	KLHRFPADLIQNMETLAIQTRDNGKISISRFQSKSAAMFRYA-----ACETTE	131		
Qy	126	GRMDASTGHKLAVIRPOPGVGLIAPPNYPNVLSGSKIAPILIGNVYMFKPPQGSYS	185		
Db	132	SSVIPSRRNYCFSFSEFVGVYALAPNNSITLEAQGLAPALAGNAIVLKSSEVTICI	191		
Qy	186	GLVIATAFAEAGLPAGVFNTITRGSEIGDYIVEHEEYVNFNFTGSPVQRIGLAKGR	245		
Db	192	GLQYGRILAMEAGFPFGILNVVTFGASVGRALVAPRGDWTFTGTGSGREIAYVGBR	251		
Qy	246	--PIMLIELGADGIVLADADLDNAAKQIVAGAYDSQORCATIKRIVVEEVADELAEX	303		
Db	252	LIPCLIELGASPNIVFDPAIDIEQAVTGMTCGIFPSNAQSCSLAGRIILOBSIYAFAR	311		
Qy	304	ISENVAKLSVDPED-NATVTPVIDNSADFIESIIVDAROKGAKELNEFKR-DGRLLTP	361		
Db	312	LTSATRALKVGPSYADDAFAVAPVSEFHHRDHIEHFVERALSGAAVALCGSGSPSGIIDK	371		
Qy	362	GLFPHVTL-----DMKLAMEEPREPLIPIRVDAAEAAVALANKSDGLOQSSVTRPQ	415		
Db	372	GAYVQPTILLEVTERSSVAHEEIFQVPAVCVMKPFQEEPLYRIANNTAAGLACGIWTADYK	431		
Qy	416	KAFIDANTLEVTGVIIINKTRGPNPFGLIGKSG---AGVGQIR-VSIEAMT	465		


```
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43455
LENGTH: 483
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-43455

Query Match      27.9%; Score 668; DB 15; Length 483;
Best Local Similarity 33.2%; Pred. No. 7.3e-45;
Matches 158; Conservative 90; Mismatches 212; Indels 16; Gaps 7;
```

```
DB      367  PFEPTLVDPVKNALVSKDETGFPLAVERFKDEAEVIAVMSNDTERGLASYFARDLARY 426
QY      418  FDIANKLEVGYTHNNKTCR-GPDNPFPLGKSGAGVCGIYISIEAMTNKSIYL 472
DB      427  FVAEQLAEYGMVGIN--TGLISNEVAPFGGIRASGIRGREGSKYGEDYIEIKYCL 480

RESULT 74
US-08-976-063C-28
Sequence 28, Application US/08976063C
Publication No. US20020182697A1
GENERAL INFORMATION:
APPLICANT: Alexander Seidhubchel; Horst Priefert; Jürgen Rabenhorst
TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF
TITLE OF INVENTION: CONFERYL ALCOHOL, CONFERYLAUDEHYDE, FERULIC ACID, VANILLIN
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: HP VECTOR
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,063C
FILING DATE: 21-NOV-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 196 49 655.1 (Germany)
FILING DATE: 29-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: Bayer 9998-CAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-063C-28

Query Match      27.9%; Score 667.5; DB 8; Length 481;
Best Local Similarity 36.4%; Pred. No. 8e-45;
Matches 169; Conservative 80; Mismatches 202; Indels 13; Gaps 6;
```

Db 250 ELGGAPFLVLDADDAVAEAAFGAYFNQGIQCMSTER-LIVTAVAAPVEKLAARKYA 308
310 KLSVGDPPFNATVY--PVIDDNGADFIESTLVDAKOKAGELMBFKRDGRLTPELFDHYT 368
Db 309 TIRAGDPNDPOSLGSLIDANAGQRIQVLDLAAAGQVGGIGGIMQPMILDQYT 368
369 LDMKLAMEEPFPIIPIIRKDAEBAVALANKSDFGLOSSVFTPRDFOKAFDIANKLEVGT 428
369 EEMRLYREESFGPVAVVLKGDDEBLRLANDSEFGLSAIFSRDYSRAMELAQRVDSGI 428
QY 429 VHINNTGRGPDNFPPLGLKSGAGVQGIKRSIEAMTNTKSIYL 472
Db 429 CHINGPTVHDEAQMFGVKSSGYSGFSRASIEHFTQLRWLTI 472

RESULT 75

US-09-750-986D-28
; Sequence 28, Application US/09750986D
; Publication No. US20030228670A1
; GENERAL INFORMATION:
; APPLICANT: Steinhuchel, Alexander
; APPLICANT: Priefert, Horst
; APPLICANT: Rabenhorst, Jurgen
; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF CONIFERYL
; TITLE OF INVENTION: ALCOHOL, CONIFERYLALDEHYDE, FERULIC ACID, VANILLIN AND
; TITLE OF INVENTION: VANILLIC ACID AND THEIR USE
; FILE REFERENCE: Bayer-9998-CAO
; CURRENT APPLICATION NUMBER: US/09/750,986D
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 196 49 655.1 GERMANY
; PRIOR FILING DATE: 1996-11-29
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Pseudomonas sp.
US-09-750-986D-28

Query Match 27.9%; Score 667.5; DB 10; Length 481;

Best Local Similarity 36.4%; Pred. No. 8e-45;

Matches 169; Conservative 80; Mismatches 202; Indels 13; Gaps 6;

QY 15 SSVNOLEISPIDDSLGVPMATREEDVHAMKAGREALPAWALTVYERAOYLHKADI 74
Db 16 SDEFTFERSSPLTGEVSVVAAASLEDDAAVAQAAPFEMALAPSERARILRAADI 75
QY 75 IERDKELATVLAKELISKAYNASVTEVVTADLIRYAAEGIRLSTADEGGRMDASTGH 134
Db 76 LE-DRSSFTAAASFTGAAGNMYGFNVYLAAGMLREA-----AMTTQIQGVIPENVP 129
QY 135 KLAVIRROQVGIVLAAPYNYPNVLSGKIAPALIGANNVMPPTQSGVSGVLAKAPA 194
Db 130 SFAMAVRQPCGVVLGIAPNNAPVILGVRVAVPLACGNTVVLKSSSELSPFTHRLIGQVLH 189
QY 195 BAGLPAGVNTITGRSE---IGDYVHEHEVNFITGSTPYGQRIKLAG--MRPIML 249
Db 190 DAGLGDGVNVVISNAPQDPAVERLIANPAVRVNFQSTIHGRITIGELISARHLKPAVL 249
QY 250 ELGGKADAGIVLADADLNAAKQIVAGAYDSGQRTAIRVLVVEEVADELAEKISENVA 309
Db 250 ELGGKAPFLVLDADDAVAEAAFGAYFNQGIQCMSTER-LIVTAVAAPVEKLAARKYA 308
QY 310 KLSVGDPPFNATVY--PVIDDNGADFIESTLVDAKOKAGELMBFKRDGRLTPELFDHYT 368
Db 309 TIRAGDPNDPOSLGSLIDANAGQRIQVLDLAAAGQVGGIGGIMQPMILDQYT 368
QY 369 LDMKLAMEEPFPIIPIIRKDAEBAVALANKSDFGLOSSVFTPRDFOKAFDIANKLEVGT 428
Db 369 EEMRLYREESFGPVAVVLKGDDEBLRLANDSEFGLSAIFSRDYSRAMELAQRVDSGI 428
QY 429 VHINNTGRGPDNFPPLGLKSGAGVQGIKRSIEAMTNTKSIYL 472

Db 429 CHINGPTVHDEAQMFGVKSSGYSGFSRASIEHFTQLRWLTI 472

Search completed: August 28, 2005, 10:40:14
Job time : 169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2005, 10:22:14 ; Search time 43 Seconds
(without alignments)
824.612 Million cell updates/sec

Title: US-09-868-195-12

Perfect score: 2393

Sequence: 1 LTKKQYVNGWKKSSVNOI.....GIRSIEMANTVKSIVLDMK 475

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 80 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Query Length	DB ID	Description
1	1808.5	75.6	474	US-09-583-110-3808	Sequence 3808, Ap
2	1692	70.7	436	US-09-107-433-3804	Sequence 3804, Ap
3	786.5	32.9	492	US-09-717-926-4	Sequence 4, Appl
4	775.5	32.4	492	US-09-634-955B-13	Sequence 13, Appl
5	733.5	30.7	526	US-09-552-991A-27302	Sequence 27302, A
6	711	29.7	487	US-09-489-039A-12633	Sequence 12633, A
7	705.5	29.5	491	US-09-585-174-3	Sequence 3, Appl
8	672	28.1	474	US-09-732-615-12	Sequence 12, Appl
9	672	28.1	474	US-10-273-051-12	Sequence 12, Appl
10	668	27.9	594	US-09-252-991A-27749	Sequence 27749, A
11	667.5	27.9	481	US-08-976-063E-28	Sequence 28, Appl
12	660	27.6	517	US-09-538-092-1047	Sequence 1047, Ap
13	659.5	27.6	508	US-08-956-171E-5241	Sequence 5241, Ap
14	659.5	27.6	508	US-08-781-986A-5241	Sequence 5241, Ap
15	658	27.5	486	US-09-328-352-4767	Sequence 4767, Ap
16	658	27.5	500	US-09-538-092-831	Sequence 831, App
17	650	27.2	497	US-09-602-777A-182	Sequence 182, App
18	650	27.2	512	US-09-538-092-1119	Sequence 1119, Ap
19	650	27.2	512	US-09-961-403-12	Sequence 12, Appl
20	645	27.0	521	US-09-321-294-2	Sequence 2, Appl
21	642.5	26.8	507	US-09-489-039A-13722	Sequence 12722, A
22	639	26.7	532	US-09-949-016-9207	Sequence 9207, Ap
23	635	26.5	498	US-09-489-039A-7569	Sequence 7569, Ap
24	632.5	26.4	505	US-09-543-681A-7151	Sequence 7151, Ap
25	632	26.4	496	US-09-543-681A-5701	Sequence 5701, Ap
26	631	26.4	529	US-09-252-991A-23785	Sequence 23785, A
27	626.5	26.2	505	US-09-328-352-5823	Sequence 5823, Ap

28	626	26.2	502	US-09-540-236-3780	Sequence 3780, Ap
29	625.5	26.1	510	US-09-252-991A-17138	Sequence 17138, A
30	624	26.1	510	US-09-134-001C-4541	Sequence 4541, Ap
31	623	26.0	493	US-09-328-352-6753	Sequence 6753, Ap
32	621	26.0	485	US-09-651-941-9	Sequence 9, Appl
33	621	26.0	485	US-09-955-597-9	Sequence 9, Appl
34	621	26.0	508	US-09-655-270A-9	Sequence 9, Appl
35	621	26.0	517	US-09-538-092-851	Sequence 851, App
36	621	26.0	520	US-09-949-016-10653	Sequence 10653, A
37	618	25.8	502	US-09-328-352-6687	Sequence 6687, Ap
38	616.5	25.8	439	US-09-949-016-9959	Sequence 9559, Ap
39	616	25.7	482	US-09-155-183-4	Sequence 4, Appl
40	616	25.7	482	US-09-733-383-4	Sequence 4, Appl
41	611.5	25.6	514	US-09-710-278-2682	Sequence 2682, Ap
42	609	25.4	497	US-08-513-841-2	Sequence 2, Appl
43	609	25.4	497	US-08-696-834-2	Sequence 2, Appl
44	609	25.4	497	US-08-942-673-2	Sequence 2, Appl
45	609	25.4	497	US-09-118-317-2	Sequence 2, Appl
46	607.5	25.4	518	US-09-134-001C-4451	Sequence 4451, Ap
47	607	25.4	501	US-09-344-882-22	Sequence 22, Appl
48	604.5	25.3	498	US-09-328-352-5847	Sequence 5847, Ap
49	602	25.2	518	US-09-976-594-287	Sequence 287, App
50	602	25.2	518	US-09-919-033-143	Sequence 143, App
51	602	25.2	553	US-09-949-016-7505	Sequence 7505, Ap
52	601	25.1	496	US-09-710-279-1030	Sequence 1030, Ap
53	601	25.1	506	US-09-134-001C-4383	Sequence 4383, Ap
54	600.5	25.1	514	US-09-489-039A-13820	Sequence 13820, A
55	599	25.0	493	US-09-538-092-1147	Sequence 1147, Ap
56	598	25.0	490	US-09-949-016-5979	Sequence 5979, Ap
57	598	25.0	490	US-09-328-352-6118	Sequence 6118, Ap
58	597.5	25.0	477	US-09-717-926-2	Sequence 2, Appl
59	595.5	24.9	477	US-09-902-546-15889	Sequence 15889, A
60	594	24.8	505	US-09-328-352-7337	Sequence 7337, Ap
61	591.5	24.7	748	US-09-252-991A-28449	Sequence 28449, A
62	590.5	24.7	500	US-09-328-352-6757	Sequence 6757, Ap
63	585	24.4	499	US-09-543-681A-4515	Sequence 4515, Ap
64	583.5	24.4	488	US-09-134-001C-4246	Sequence 4246, Ap
65	581.5	24.3	492	US-09-489-039A-9023	Sequence 9023, Ap
66	578	24.2	502	US-09-489-039A-13971	Sequence 13971, A
67	577.5	24.1	481	US-09-328-352-4475	Sequence 4475, Ap
68	568.5	23.8	538	US-09-344-882-20	Sequence 20, Appl
69	562.5	23.5	486	US-09-716-865-6	Sequence 6, Appl
70	562	23.5	475	US-09-252-991A-32298	Sequence 32298, A
71	561	23.4	642	US-09-252-991A-25398	Sequence 25398, A
72	559.5	23.4	534	US-09-344-882-24	Sequence 24, Appl
73	555	23.2	535	US-09-538-092-1173	Sequence 1173, Ap
74	555	23.2	564	US-09-328-352-7905	Sequence 7905, Ap
75	554.5	23.2	468	US-09-351-150A-9	Sequence 9, Appl
76	553.5	23.1	493	US-09-134-001C-4388	Sequence 4388, Ap
77	552	23.1	477	US-09-328-352-6549	Sequence 6549, Ap
78	551.5	23.0	479	US-09-710-278-2956	Sequence 2956, Ap
79	550.5	23.0	492	US-09-252-991A-26724	Sequence 26724, A
80	539.5	22.5	498	US-09-648-004-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-583-110-3808
Sequence 3808, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
FILE REFERENCE: PATHO0-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12

PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO: 3808
LENGTH: 474
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3808

Query Match 75.6%; Score 1808.5; DB 4; Length 474;
Best Local Similarity 73.7%; Pred. No. 4.7e-161;
Matches 350; Conservative 59; Mismatches 65; Indels 1; Gaps 1;

QY 1 LTRKQYNYNGEMKSSVNOIETLSPIDSSLGFPVPMREBYDHAMKAGREALPMAALT 60
DB 1 LTRKQYNYNGEMKSSVNOIETLSPIDSSLGFPVPMREBYDHAMKAGREALPMAALT 59
QY 61 YVERAQLYHKAADIIRDKKEIATYLAKESKAYNASVTEVVRTADLIRYAAEGIRLST 120
DB 60 ATERAAYLHKTALIERDKEKIGTILAKEVAKIKALIGEVVRTADLIRYAAEGIRLST 119
QY 121 SADEGGKMDASTGHKLAVRRQPGIVLAIAPVNPNVSGSKIPALIGNVNFKPPT 180
DB 120 QMEGGGFEAASKNKLAVRRPVGIVLAIAFPNPNVLSASKIAPALIAGVNFKPPT 179
QY 181 QGSVGLVLAKAFAPAGLPAGVFNITTRGSEIGDIYVEHEEVNFINFGSTPVGQRIK 240
DB 180 QGSISGLLAKAFEPAGIPAGVFNITTRGSEIGDIYIHKVNFINTFGSTPIGERICR 239
QY 241 IAGMRPIMELGKQDAGIVLADADLNAAKQIVAGAYDYSQRCIAIKRVLVEEVADEL 300
DB 240 IAGMRPIMELGKQDAGIVLADADLNAAKQIVAGAFYSQRCIAIKRVILVESVADKL 299
QY 301 AEKISENVAKLSVGPDPNATVTPVIDNSADPISLVYDAROKAKELNEKRGRLIT 360
DB 300 ATLLEBVSCLTVGDPFNADITPVINDNSADPIMGLIDAOEKEQALTPIKRENNILW 359
QY 361 PGLFDHVTLDKLAEEFPFPIIPRIYKDAEAVAIANKSDFGLQSSVFTRDFOKAFDI 420
DB 360 PGLFDHVTLDKLAEEFPFPIIPRIYKDAEAVAIANKSDFGLQSSVFTRDFOKAFDI 419
QY 421 ANKLEVGTVHINNKTGRGPDNFPFLGKSGAGVQIGIRYSIEMTNVKSIVLDMK 475
DB 420 AEKLEVGTVHINNKTGRGPDNFPFLGKSGAGVQIGIRYSIEMTNVKSIVPDVK 474

RESULT 2
US-09-107-433-3804
Sequence 3804, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID and AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131

FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Atinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3804:
SEQUENCE CHARACTERISTICS:
LENGTH: 436 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...436
SEQUENCE DESCRIPTION: SEQ ID NO: 3804:
US-09-107-433-3804

Query Match 70.7%; Score 1692; DB 4; Length 436;
Best Local Similarity 74.9%; Pred. No. 3.6e-150;
Matches 326; Conservative 53; Mismatches 56; Indels 0; Gaps 0;

QY 41 EVDHAKAGREALPMAALTVERAQLYHKAADIIRDKKEIATYLAKESKAYNASVTE 100
DB 2 EADEGQARALPMAALPMAALSAIERAAYLHKTALIERDKEKIGTILAKIKKALIG 61
QY 101 VVRTADLIRYAAEGIRLSTADSGKMDASTGHKLAVRRQPGIVLAIAPVNPNVLS 160
DB 62 VVRTADLIRYAAEGIRLSTGQMEGGGFEAASKNKLAVRRPVGIVLAIAFPNPNVLS 121
QY 161 GSKIPALIGNVNFKPPTQGSVGLVLAKAFAPAGLPAGVFNITTRGSEIGDIYEH 220
DB 122 ASKIPALIAGVNFKPPTQGSISGLLAKAFEPAGIPAGVFNITTRGSEIGDIYIEH 181
QY 221 EEVNFINFGSTPVGQRIKLAGMRPIMELGKQDAGIVLADADLNAAKQIVAGAYDYS 280
DB 182 KEVNFINTFGSTPIGERIGRLAGMRPIMELGKQDAGIVLADADLNAAKQIVAGAFYS 241
QY 281 GORCTAIRVLVEEVADELAEKISENVAKLSVGPDPNATVTPVIDNSADPISLVYD 340
DB 242 GORCTAIRVLVEEVADELAEKISENVAKLSVGPDPNADITPVINDNSADPIMGLIED 301
QY 341 AROKAGELNFKRGRLITPGLFDHVTLDKLAEEFPFPIIPRIYKDAEAVAIANK 400
DB 302 AOEKAQALTPIKRENNILWVLPDQYTKDMKVAEEFPFPIIPRIYKDAEAVAIANK 361
QY 401 SDFGLQSSVFTRDFOKAFDIANKLEVGTVHINNKTGRGPDNFPFLGKSGAGVQIGIRYS 460
DB 362 SEFGLQSSVFTNDFOKAFDIANKLEVGTVHINNKTGRGPDNFPFLGKSGAGVQIGIRYS 421
QY 461 IEAMTNVKSIVLDMK 475
DB 422 IEAMTNVKSIVPDVK 436

RESULT 3
US-09-717-926-4
Sequence 4, Application US/09717926
Patent No. 6569657
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
APPLICANT: Cook, William J.
TITLE OF INVENTION: 32140, A No. 6569657el Human Aldehyde
TITLE OF INVENTION: Dehydrogenase
FILE REFERENCE: 35800/205243
CURRENT APPLICATION NUMBER: US/09/717, 926


```
/ CURRENT FILING DATE: 2000-11-21
/ PRIOR APPLICATION NUMBER: US 60/214,707
/ PRIOR FILING DATE: 2000-06-27
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 492
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Aldehyde dehydrogenase Pfam consensus sequence
US-09-717-926-4

Query Match      32.9%; Score 786.5; DB 4; Length 492;
Best Local Similarity 38.6%; Pred. No. 4.5e-65;
Matches 193; Conservative 92; Mismatches 166; Indels 49; Gaps 16;

QY 12 EWKSSVN--QIEILSPIDSS--LGFVPAMTREVVDHAKAGREAL---PAMALTYVERA 65
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
1 EWVDASAGKTFEVNPNANGVEVIGRVPEATADVDAAVAAKAEAFKSGPMWAKVPASERA 60

QY 66 QYIHKAAADIERDKKEIATVLAKEISK--AYNASVTEVVRTADLRYYA-----EEG 115
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
61 RILRLKADLIERDEDLALETLDLCKPLAEAKGTEVGRALDEIRYAGMARKLMGERR 120

QY 116 I--RLSTADGCKDASTGHKLAVIRROPVGIYLAIPYNPVNLSSGIAPALIGNV 173
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
121 VLPSLATDGE-----ELNYYRREPGLGVGVISPMNPLALMLKLAIPALAAQNT 170

QY 174 VMEKPTQGSVSGVLAKAPAEAG--LPAGVFNTITGRGSEIGDYVEHEEVNFINFTG 230
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
171 VVLRKSEQPTPLALLAEILIEEGANNLPKGVNVVPGGAEVQALLSHPIDIKISFTG 230

QY 231 STPVGQRIGKLA---GMRPIMELGCKDAGIYLDADLDNAAKOIYAGAYDSGQRCTAI 287
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
231 STEVGLIMEAAAKNLKVTIELGKSPVIYFDADLDKAVRIVFGAFNAGGVCIAP 290

QY 288 KAVLVEEYADLAEKISENVAKLS--VGDPFDMAT--VTPVIDNSADPISLAVDAROK 344
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
291 SKLHVHESIYDFEVEKLERVKKLKIDPLDSNTNIGPLISEQGFDPVLSYIEDGKEE 350

QY 345 GAK-----ELNEFKRDGRLLTPGLFDHVTLDMLAMEBPPGRLPIIRYKDAEEAVA 396
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
351 GAKVLCGGRDSKELYGGGYVQPTIFDVTIDMKIMEELFGVPLPIIKFDLDEALE 410

QY 397 IANKSDPGLQSSVFTRD--FOKAFDIANKLEVTGTHINNTGKRPD--NFPPLGLK--GSGAG 453
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
411 LANDREYGLAAVYFTDILARAFVAKALEAGIWNVDVCVAHAEPQLPFGGVKSSGIG 470

QY 454 VO-GIRYSIEAMTNVSIYL 472
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
471 REHGKYGLEEYTEIKVTI 490

Db

RESULT 4
US-09-634-955B-13
/ Sequence 13, Application US/09634955B
/ Patent No. 6511834
/ GENERAL INFORMATION:
/ APPLICANT: Meyers, Rachel
/ TITLE OF INVENTION: 32142, 21481, 25964, 21686, NOVEL HUMAN DEHYDROGENASE
/ FILE REFERENCE: MNI-134
/ CURRENT APPLICATION NUMBER: US/09/634,955B
/ CURRENT FILING DATE: 2000-08-08
/ PRIOR APPLICATION NUMBER: 60/192,002
/ PRIOR FILING DATE: 2000-03-24
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: FastSeq for Windows Version 4.
/ SEQ ID NO 13
/ LENGTH: 492
/ TYPE: PRT
```

```
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Aldehyde dehydrogenase family domain
US-09-634-955B-13

Query Match      32.4%; Score 775.5; DB 4; Length 492;
Best Local Similarity 38.4%; Pred. No. 4.9e-64;
Matches 192; Conservative 92; Mismatches 167; Indels 49; Gaps 16;

QY 12 EWKSSVN--QIEILSPIDSS--LGFVPAMTREVVDHAKAGREAL---PAMALTYVERA 65
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
1 EWVDASAGKTFEVNPNANGVEVIGRVPEATADVDAAVAAKAEAFKSGPMWAKVPASERA 60

QY 66 QYIHKAAADIERDKKEIATVLAKEISK--AYNASVTEVVRTADLRYYA-----EEG 115
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
61 RILRLKADLIERDEDLALETLDLCKPLAEAKGTEVGRALDEIRYAGMARKLMGERR 120

QY 116 I--RLSTADGCKDASTGHKLAVIRROPVGIYLAIPYNPVNLSSGIAPALIGNV 173
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
121 VLPSLATDGE-----ELNYYRREPGLGVGVISPMNPLALMLKLAIPALAAQNT 170

QY 174 VMEKPTQGSVSGVLAKAPAEAG--LPAGVFNTITGRGSEIGDYVEHEEVNFINFTG 230
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
171 VVLRKSEQPTPLALLAEILIEEGANNLPKGVNVVPGGAEVQALLSHPIDIKISFTG 230

QY 231 STPVGQRIGKLA---GMRPIMELGCKDAGIYLDADLDNAAKOIYAGAYDSGQRCTAI 287
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
231 STEVGLIMEAAAKNLKVTIELGKSPVIYFDADLDKAVRIVFGAFNAGGVCIAP 290

QY 288 KAVLVEEYADLAEKISENVAKLS--VGDPFDMAT--VTPVIDNSADPISLAVDAROK 344
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
291 SKLHVHESIYDFEVEKLERVKKLKIDPLDSNTNIGPLISEQGFDPVLSYIEDGKEE 350

QY 345 GAK-----ELNEFKRDGRLLTPGLFDHVTLDMLAMEBPPGRLPIIRYKDAEEAVA 396
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
351 GAKVLCGGRDSKELYGGGYVQPTIFDVTIDMKIMEELFGVPLPIIKFDLDEALE 410

QY 397 IANKSDPGLQSSVFTRD--FOKAFDIANKLEVTGTHINNTGKRPD--NFPPLGL--KSGAG 453
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
411 LANDREYGLAAVYFTDILARAFVAKALEAGIWNVDVCVAHAEPQLPFGGVHOSGIG 470

QY 454 VO-GIRYSIEAMTNVSIYL 472
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
471 REHGKYGLEEYTEIKVTI 490

Db

RESULT 5
US-09-252-991A-27302
/ Sequence 27302, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 27302
/ LENGTH: 526
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27302

Query Match      30.7%; Score 733.5; DB 4; Length 526;
Best Local Similarity 36.8%; Pred. No. 4.8e-60;
Matches 172; Conservative 87; Mismatches 186; Indels 23; Gaps 7;

QY 11 GEMK--SSVNOIEILSPIDSSLGFVPAMTREVVDHAKAGREALPAMALTYVERAQL 68
```

```
Db      53 GEMRGRGRRLKVNPFDDGSLLEBQADRDLDAAVAKAEVOPAWAALPPSARAATL 112
Qy      69 HKAADIIEBDEKKEITVLAKEISKAYNASVTEVVRADLIRYAEEGISTSTA-----DE 124
Db      113 YKAVEFPDRRHEIYDWITRESG-----TRIKAEIENGARAILTESASFPARVH 163
Qy      125 GGRMDASTGHLKLAIVRQPVGIVLAIAPYNYPVNLSGSKIAPALIGNVVMEKPPPTGVS 184
Db      164 GRIVSDVFGKESRYRSALIGVGVISFPWNPFLHLTGRIAPALALGNVAVVKPASPDPV 223
Qy      185 -SGVLAKAPAEAGLPAGVFNTITGSGEIGDIYVHEHEVNFINFTGSPVQRIKLAG 243
Db      224 CGGLIARIFEBAAGPAGLFSYVVGSGSEIGDAFVEHPVPGIYTFGTSGTPVGRNIGRIAS 283
Qy      244 ----NRPIMLELGGKDAIGVLADADLDNNAKQIVAGAYVSGORCAIKRIVVVEEVD 299
Db      284 GGAHLKHVALLELGSNPFVVLGDADLEQVNAAPFEGKFLHGGQICMAINCITIEBSDLDA 343
Qy      300 LAEKISENVAKLSVGP--FDNATVTPVIDNSADFIESLVVDAROKAKEINEFKRDR 357
Db      344 FAFRFRVRYKGLRVGDPQADTA--VGPVYNAQLBGLKIKILARQEGAKPLYEGSVDOQ 402
Qy      358 LITPGLFDHVTIDMKLAMEEPPGPIPIRVKDAEBAVAIAKSDPGLGSSVFTTRDPQKA 417
Db      403 ILAPHVFEVITATMEIARDEIFGPIVGLRLARDEAHLIELANASEYGLSSAVFSRDLERA 462
Qy      418 FDIANKLEGVTVHINKTG-RGPDNPFPLGLKSGAGVQGIKRISEAMT 465
Db      463 VFAFQQLRAGMTHVDIPVNDANAPFGEKNSGLGRFNGDWAIEEFT 510
```

RESULT 6
US-09-489-039A-12633
Sequence 12633, Application US/09489039A

Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709 2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 12633
LENGTH: 487
TYPE: PRF
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12633

Query Match 29.7%; Score 711; DB 4; Length 487;
Best Local Similarity 34.1%; Pred. No. 5, Se-58;
Matches 163; Conservative 90; Mismatches 209; Indels 16; Gaps 7;
Qy 6 QNYVNGEMKSVN--QIETLSPIIDSSLGFPVPMTEEEVDHAKKAGREALPAAATLYVE 63
Db 17 QAMIDGQMDANGDVIAVTNPANGEOGQSVPMKMGDREIRAEANRALPAMRALITAAE 76
Qy 64 RAQYIHKAAADIIEBDEKKEIATVLAKEISKAYNASVTEVVRADLIRYAEEGISTSTA 123
Db 77 RANIIIRRMFDLMNEMDDLARIMLTLEQKPLAEAKESISYAASFIEWFAEEGKRIVGDTI 136
Qy 124 EGGKMDASTGHLKLAIVRQPVGIVLAIAPYNYPVNLSGSKIAPALIGNVVMEKPPPTGS 183
Db 137 PEHQAD----KRLVTKPIKPIGIVTAITWNPAAAITRKAGPALAAGCTWALKPASQTP 191
Qy 184 VSGVLAKAPAEAGLPAGVFNTITGSGEIGDIYVHEHEVNFINFTGSPVQRIKLAG 243
Db 192 FSAALALAEIARAGIPAGVFNVTGSGAGVGEILTSNPLVRKISFGSIEIRQLMEQCA 251
Qy 244 --MRPIMLELGGKDAIGVLADADLDNNAKQIVAGAYVSGORCTAIKRLVVEEVADELA 301

```
Db      252 KDIKVSLELGGNAPFIFDDADLDKAVEGALASKFRNAGATCCANLTYQDGVYDRFA 311
Qy      302 EKISENVAKLSVGP--DNATVTPVIDNSADFIESLVVDAROKAKEINEFK--RDGR 357
Db      312 EKLQQAVERKLIQDLDQGVTTGPIIDKAKAYKEHIAADIAKAKAVYTGKPHALGQ 371
Qy      358 LITPGLFDHVTIDMKLAMEEPPGPIPIRVKDAEBAVAIAKSDPGLGSSVFTTRDPQKA 417
Db      372 FPGPIIIVNVDGSAKVAKEETFGPLAPFRFKDEADVIAQANDTEFGIAAFVARDLSRV 431
Qy      418 FDIANKLEGVTVHINKTG-RGPDNPFPLGLKSGAGVQGIKRISEAMTNYKSIYLD 474
Db      432 FRVGRALRYGIIGIN--TGIISTEVAAPGVKASGLGREGSKYGIEDIYELIKMCTGL 487
```

RESULT 7
US-09-585-174-3
Sequence 3, Application US/09585174

Patent No. 6586229
GENERAL INFORMATION:
APPLICANT: Ben-Bassat, Arye
APPLICANT: Catermole, Monica
APPLICANT: Gatenby, Anthony A.
APPLICANT: Gibson, Katherine J.
APPLICANT: Ramos-Gonzalez, Isabel
APPLICANT: Ramos, Juan
TITLE OF INVENTION: Method for the production of p-Hydroxybenzoate in Species of
FILE REFERENCE: BCI018 US NA
CURRENT APPLICATION NUMBER: US/09/585,174
CURRENT FILING DATE: 2000-06-01
NUMBER OF SEQ ID NOS: 112
SOFTWARE: Microsoft Office 97
SEQ ID NO 3
LENGTH: 491
TYPE: PRF
ORGANISM: Pseudomonas mendocina KR-1
US-09-585-174-3

Query Match 29.5%; Score 705.5; DB 4; Length 491;
Best Local Similarity 34.5%; Pred. No. 1, Se-57;
Matches 167; Conservative 94; Mismatches 198; Indels 25; Gaps 8;
Qy 5 YONY-----VNGEWMK--SSVNOIEILSPIDSSLGFPVPMTEEEVDHAKKAGREALPAA 58
Db 8 YONLEQPLAGOMRAGSSGRPLEVFPDYNDELRLIALASREDDAAVRAKADSGREWAT 67
Qy 59 LTVERRAQYLKKAADIIEBDEKKEIATVLAKEISKAYNASVTEVVRADLIRYAEEGIRL 118
Db 68 TAPARARAVLEAVVIFPERREIYDWITRESG-----TRIKAEIENGARAILTESASFP 118
Qy 119 STSADG--GKMDAST-GHKLAVTRQPVGIVLAIAPYNYPVNLSGSKIAPALIGNAV 174
Db 119 ESASIPNKHVGRILINSNGKESRYRAPLVGIIGISFPWNPFLHLTGRIAPALALGN 178
Qy 175 MEKPPPTGVS--GIVLAKAPAEAGLPAGVFNTITGSGEIGDIYVHEHEVNFINFTGSP 233
Db 179 VVKPASTPPIITGGLIARIFEBAAGLPAGVLSVVGSGEIGDAFVEHPVPGIYTFGTSGT 238
Qy 234 VGORIGKLAG----NRPIMLELGGKDAIGVLADADLDNNAKQIVAGAYVSGORCTAIK 289
Db 239 VGRNIGRIASGEHLKHVALLELGSNPFVVLGDADLEQVNAAPFEGKFLHGGQICMAINR 298
Qy 290 VLVVEEVADELAEKISENVAKLSVGDPPDNATV--PVTDNSADFIESLVVDAROKAKE 348
Db 299 IIVEQPLLEDFTRFRVFRVYKALPYGDPBKPGTVVGPVYNAQLAGLKEKIKATAKEGATL 358
Qy 349 LNEFKRDRRLITPGLFDHVTIDMKLAMEEPPGPIPIRVKDAEBAVAIAKSDPGLGSS 408
Db 359 LUGSEPQGNVMPHFVGNVTADMEIARDEIFGPIVGLQISARDAEHLIELANASEYGLSSA 418

;; PRIOR APPLICATION NUMBER: US 60/094,190
 ;; PRIOR FILING DATE: 1998-07-27
 ;; NUMBER OF SEQ ID NOS: 33142
 ;; SEQ ID NO 27749
 ;; LENGTH: 594
 ;; TYPE: PRT
 ;; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-27749

Query Match 27.9%; Score 668; DB 4; Length 594;
 Best Local Similarity 33.2%; Pred. No. 8.2e-54;
 Matches 158; Conservative 90; Mismatches 212; Indels 16; Gaps 7;

QY 6 QNYVNGEKSSVNV--QIILSPIDSSLSGFVPAMTREEDVHAMKAGREALPMAALTYVE 63
 DB 123 QAYVGDVADNDGQTIKVNNDPATGEIIGSVPMGAETRRALIEADAKLAPMALTLAKE 182
 QY 64 RAQYIHKADIIERKEEIAIYLAKEISKAYNASTEYVRTDLIRYAEGRISTAD 123
 DB 183 RANKLRWFEDMIENQDILRLMTIEQKPLAEAKGEIAYAAFLPEWFEERKRIYDPTI 242
 QY 124 EGGKMDASTGKHLAVIRPQVIGVLAIAFYNPVNLSSGSKIAPALIGVNVMEKPTQGS 183
 DB 243 PGHQPD-----KRIYIKQPIGVTAIITPMNPPSAMITRKAPALAGCTMVLKASQTP 297
 QY 194 VSGVLAKAFAPAGLPAAGVNTITGRSGEIGDYIYEHEEVNFINTGSTPVGQRIK--L 241
 DB 298 YSALALAEIAPAGIPKGVFSVVTGSAGVGGELNSPIVRKLTFGTSGTEIGRLMAEGA 357
 QY 242 AGMRIMLELGKDGIVLADADLNNAAQIYAGAYDSGQRTAIKRLVVEVVADELA 301
 DB 358 QDIKEVSELEGGNAPFIYFDADLDAVAEGALISKRNNQCTCVCANRLYVODGYDAFV 417
 QY 302 EKISNVAKLSVGDFPD-NATVTPVLDNSADPISLVVDARQKAKELNEFK---RDR 357
 DB 418 DKLKAQVAKLNGKLEAGVTTGPLIDAKAVAKVEEHIDAVSKGAKVSGKPHALGCT 477
 QY 358 LITPGLFPHVTLDMKLAMEBPPIPIIRVDAEBAVAIAKSDPGLQSSVFTDFOKA 417
 DB 478 FPEPTLVDPFGNALVSKDETFGPIAPVFRFOBAEVLAMSDTEFGLASFYADLARI 537
 QY 418 FDIANKLEVGTTHINNKTR-GPDNPPPLGLKSGAGVQIGIRYSIEAMNVKSIYL 472
 DB 538 FRVAQLEYGWGIN--TGLISNEVAPFGIGIKASGLGRGSKYGIEDYLEIKYLCL 591

RESULT 11
 US-08-976-063E-28
 ; Sequence 28, Application US/08976063E
 ; Patent No. 6524831
 ; GENERAL INFORMATION:
 ; APPLICANT: Steinhuchel, Alexander
 ; APPLICANT: Priefert, Horst
 ; APPLICANT: Rabenhorst, Urogen
 ; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF CONIFERYL
 ; TITLE OF INVENTION: ALCOHOL, CONIFERYLALDEHYDE, FERULIC ACID, VANILLIN AND
 ; TITLE OF INVENTION: VANILLIC ACID AND THEIR USE
 ; FILE REFERENCE: Bayer-9998-CNO
 ; CURRENT APPLICATION NUMBER: US/08/976,063E
 ; CURRENT FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 196 49 655.1 GERMANY
 ; PRIOR FILING DATE: 1996-11-29
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 28
 ; LENGTH: 481
 ; TYPE: PRT
 ; ORGANISM: not required under old rule
 US-08-976-063E-28

Query Match 27.9%; Score 667.5; DB 4; Length 481;
 Best Local Similarity 36.4%; Pred. No. 6.5e-54;
 Matches 169; Conservative 80; Mismatches 202; Indels 13; Gaps 6;

QY 15 SSVNOIEILSPIDSSLSGFVPAMTREEDVHAMKAGREALPMAALTYVERAQLYHKADI 74
 DB 16 SDBRTFERSPLTGEVVSRAVAASLEDDADAAVAAGAFPMALALPESERRALIRADL 75
 QY 75 IERDKEEIAIYLAKEISKAYNASTEYVRTDLIRYAEGRISTADDEGKMDASTGH 134
 DB 76 LE-DRSSEFTAAASFTGAAGWYGFNVYLAAGMREAA-----AMTTOIGDVIIPSNVPG 129
 QY 135 KLAIVRRPQVIGVLAIAFYNPVNLSSGSKIAPALIGVNVMEKPTQGSGLVLAKAPA 194
 DB 130 SPMAVNRQPCGVVIGIAFMNAPVILGVRAVMPALCGTIVLKSSSELPFTHRLIGQVLH 189
 QY 195 BAGLPAGVNTITGRSGE---IGDYIYEHEEVNFINTGSTPVGQRIKLAG--MRPIML 249
 DB 190 DAGIGDGVVNNIISNAPQAPAVVERLIANPAVRNRNFGSTHYGRIIGELSRHLKPAVL 249
 QY 250 ELGKDGIVLADADLNNAAQIYAGAYDSGQRTAIKRLVVEVVADELAKEISENVA 309
 DB 250 ELGGKAPPLVLDADLDAVAEAAAFGAYFNOGOICMSTER-LIYVAVADAFVEKLAARVA 308
 QY 310 KLSVGDPPDNATVT--PVLDNSADPISLVVDARQKAKELNEFKRDGRLITPGLFDHYT 368
 DB 309 TLRAGDPPDPOSVGLSLIDANAGORIQVLVDLAKAKARQYVGGGLDGSIMQPHLLDQYT 368
 QY 369 LDMKLAMEBPPIPIIRVDAEBAVAIAKSDPGLQSSVFTDFOKAPDIANKLEVGT 428
 DB 369 EBMRLYRESGPAVAVLRGPDDELLRLANDSEFGSAALFSRVSAMELAQVDSGI 428
 QY 429 VHINNKTRGPDNPPPLGLKSGAGVQIGIRYSIEAMNVKSIYL 472
 DB 429 CHINPVTVDHDAQWPFGGVSKSGYGSFSGRASIEHFTQLRWLTI 472

RESULT 12
 US-09-538-092-1047
 ; Sequence 1047, Application US/09538092
 ; Patent No. 6753314
 ; GENERAL INFORMATION:
 ; APPLICANT: Giot, Loic
 ; APPLICANT: Mansfield, Traci A.
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
 ; FILE REFERENCE: 15966-542
 ; CURRENT APPLICATION NUMBER: US/09/538,092
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/127,352
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178,965
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 1387
 ; SOFTWARE: CuraPatSeqFormatter Version 0.9
 ; SEQ ID NO 1047
 ; LENGTH: 517
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (0)-(0)
 ; OTHER INFORMATION: Polypeptide Accession Number P30837
 US-09-538-092-1047

Query Match 27.6%; Score 660; DB 4; Length 517;
 Best Local Similarity 34.8%; Pred. No. 3.7e-53;
 Matches 170; Conservative 80; Mismatches 201; Indels 38; Gaps 11;

QY 6 QNYVNGEKSSVNVQ--IILSPIDSSLSGFVPAMTREEDVHAMKAGREAL--PMAALTY 60
 DB 38 QLFINWQDQAVSKTPTVNTPTGEVIGYVAEGDRAVDRAVAARAARFGLGSPWRMD 97
 QY 61 YVERAQLYHKADIIERKEEIAIY----LAKESKANASVTEYVRTDLIRYAEGRIST 116
 DB 98 ASERGRLLNLADLVERDRVYLAILETLNKGKPPQESYALDLDEVIK---VYRYFA---- 150


```
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Benson, Bob
/ REGISTRATION NUMBER: 30,446
/ REFERENCE/DOCKET NUMBER: PB248PP
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 5241:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 508 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-781-986A-5241

Query Match      27.6%; Score 659.5; DB 4; Length 508;
Best Local Similarity 34.0%; Pred. No. 4e-53;
Matches 166; Conservative 99; Mismatches 198; Indels 25; Gaps 10;

QY 1 LTKSEYQNYNGEM--KSSVNOJEIISPIDSSLGFPVPMTRREVDHAKAGREALPAMAA 58
DB 22 IAEHTGLFNGEFVKSSDETELEVTPATGETLSHTTRAKDQVDHVAQAEAFESWLSL 81
QY 59 LTVERRAQLYHKAADIIEERKEEIAVLAKESKAYNASVTEVTRADLIRYAAEE-GIR 117
DB 82 TSKSRAQWLRIQGLKMAQKIMIEFLNNGKPIREF-----TADIPRAHHFHF 135
QY 118 LSTSDGEGKMDASTGHKLAIVTRROPVGIYLAIAFYNYPNVLSGSKIAPALIGSNVMEK 177
DB 136 ASVIEETEGTVNDIDKDTMSIVRHEBIGVVGAVVAMNPFMLAAMKIAIPALIAAGNTIYIQ 195
QY 178 PRTQGSVSGVLAKAPAEAGLPAGVFNTITGRSGEIGDIYVEHEVNFINFSTPVGQR 237
DB 196 PSSPPLSLAEVAKIFOEV-LPKGVNNTITGSGESGNALFNHDDVDKLSFGSTDVQ 254
QY 238 ICKLAG-MRPIMLEIGKDGAGIVLADADLDNAKOIVAGAYDSGORCTAIKRVLYVEE 295
DB 255 VAAEAAKHLVPATLEIGKGSANIIDDANLDLAVESIQGLIFNGEVSASGRLLVHEK 314
QY 296 VADSLAEKISENVAKLSVGDPRDNAT-VTPVIDNSADFTESLVDAKQKAEI--NE 351
DB 315 IYDQLVPRLOEAFSNIKVGNPDDEATQMGSGTGKQQLDKIQSYIDAAKESDAQIILAGHR 374
QY 352 FRDGRLLTPGLFDHYTL-----DMKLAEEEPFGIILPIIVKAAEEVAIALNKSDFGL 405
DB 375 LTENG-LDKGFEFEPTLIAVDNHHKLAQEEIFGVLTIVIKVDQDEALIDANDSEYGL 432
QY 406 OSSVETRDFOKAFDIANKLEVGTVHINNKTGRGPDNFPPLGLKSGAGVQGIYSIEAMT 465
DB 433 AGGVSONITRALNLTAKAVRTGRIMWT-YNQVPBAGPFGYKKSIGSEYTKGALSNTQ 491
QY 466 NYKSLIVLD 473
DB 492 QVKNYIID 499

RESULT 15
US-09-328-352-4767
/ Sequence 4767, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 4767
/ LENGTH: 486
```

```
/ TYPE: PRT
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-4767

Query Match      27.5%; Score 658; DB 4; Length 486;
Best Local Similarity 32.2%; Pred. No. 5.2e-53;
Matches 154; Conservative 92; Mismatches 219; Indels 14; Gaps 6;

QY 4 EYQYVNGEMKSSVNO--IEIISPIDSSLGFPVPMTRREVDHAKAGREALPAMAAITLV 61
DB 14 QQQAYINQWMLAAGSNAPVPSNPATGTEIGTIPRMGALEATQVAEAYTLQSKALTA 73
QY 62 YERRAQLYHKAADIIEERKEEIAVLAKESKAYNASVTEVTRADLIRYAAEEGIRLSTS 121
DB 74 QNRADIIILAMHKLVLIDHDELALIMTIEQKPLAEAKGEVRYAASFQWFAEEGRKI--- 130
QY 122 ADEGKMDASTGHKLAIVTRROPVGIYLAIAFYNYPNVLSGSKIAPALIGSNVMEKPTQ 181
DB 131 --YGDVITPVNNQORFIISKEPVGVAITPMNPFIAITRKAPALAAAGCTVYIKPANE 188
QY 182 GSVSGVLAKAPAEAGLPAGVFNTITGRSGEIGDIYVEHEVNFINFSTPVGQRIGKL 241
DB 189 TPYCALAIKAEKAGIYAGVINYVTGKSQIEGVSFTHSEKVKLTFTGSTPVGRLMQ 248
QY 242 AG-MRPIMLEIGKDGAGIVLADADLDNAKOIVAGAYDSGORCTAIKRVLYVEEVADE 299
DB 249 CSSTIKKLALIEGNAPIVFPDDADLDKAVOGAIFAKFRNAGQVCVCANRIYVDNIYQA 308
QY 300 LAEKISENVAKLSVGDPR-DNATVTPVIDNSADFTESLVDAKQKAEIENFPRDG-- 356
DB 309 FAEKRVQEVQKFGVNGLEDVQIGPLINEKAVLKAQQLIDDAVSKAKIACGGKQHALG 368
QY 357 -RLTPGLFDHYTLDMKLAEEEPFGIILPIIVKDAEBAVAILNKSDFGLQSVETRDFO 415
DB 369 QTFPEPSVLTVNDRMEIVQEEIFGPVAPLIRTFDEADVAVQAQNDITIGLAAYISENIS 428
QY 416 KAFDIANKLEVGTVHINNKTGRGPDNFPPLGLKSGAGVQGIYSIEAMTVKSLIVDM 474
DB 429 RLMRVSEQLIEGVGMGM-NATAISNEVVPFGVKGSGVREGSKYGLBEFMTIKYWCGL 486

RESULT 16
US-09-538-092-831
/ Sequence 831, Application US/09538092
/ Patent No. 6753314
/ GENERAL INFORMATION:
/ APPLICANT: Giot, Loic
/ APPLICANT: Mansfield, Traci A.
/ TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
/ FILE REFERENCE: 15966-542
/ CURRENT APPLICATION NUMBER: US/09/538,092
/ CURRENT FILING DATE: 2000-03-29
/ PRIOR FILING DATE: 1999-04-01
/ PRIOR APPLICATION NUMBER: 60/127,352
/ PRIOR FILING DATE: 1999-04-01
/ PRIOR APPLICATION NUMBER: 60/178,965
/ NUMBER OF SEQ ID NOS: 1387
/ SOFTWARE: CuratPatSeqFormatter Version 0.9
/ SEQ ID NO 831
/ LENGTH: 500
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (0)...(0)
/ OTHER INFORMATION: Polypeptide Accession Number P00352
US-09-538-092-831

Query Match      27.5%; Score 658; DB 4; Length 500;
Best Local Similarity 34.1%; Pred. No. 5.4e-53;
Matches 163; Conservative 85; Mismatches 210; Indels 20; Gaps 8;

QY 8 YVNGEMKSSVN--QIEIISPIDSSLGFPVPMTRREVDHAKAGREAL--PAMAAITLV 62
```

```
Db      23 FINNEHDSVSGKKFPEVNPATEEELQVEGEKEVDKAKARQAFOIGSWMRTMDAS 82
      63 ERAOYLHKAAADIIEERDEEIAIVLAKEISKAY-NASVTEVTRADILRYAAEEGIRLSTS 121
      83 ERERLLYKLAIDLIERDRLLATWESNMGKLYSNAYLADLACIKTLRYCAGWADKIQ-- 140
Qy      122 ADGGCKMDASTGKKAIVIRQPIGIVLAIPYVYPNVLSSKTAAPALIGNVMPFPPTQ 181
      141 ---GRTIPIDGNFPTTYTRHEPIGVCGQIIPMNPFLMWLWKIPALISGNTVAVVPAEQ 196
Qy      182 GSVGVLAKAFPAEAGLPAGVFNTITGRGSEIDYIYEHEVEVNFINTGSTPVGORIGKL 241
      197 TPLTALHVASLIEAGEFPEPVNAIVPEYGTAGAAALISSHMDIKVAFSTGEVGLIKEXA 256
Qy      242 AG--MRPIMLELGGKAGIVLADADLDNNAKOIVAGAYDYSQORCTAIRVLVEEVD 298
      257 AGKSNLKRVTLLELGGKSPCIVLADADLDNAVEFAHHGVFHQGCCIAAGRIFVESIYD 316
Qy      299 ELAEKISENVAKLSVGDPF-DNATVTPVIDNSADPIESLVDAKOKAK--ELAEFKR 354
      317 EFVRREVERAKKYLINPILPVGTOGPQIDKEQYDKILDLESKKEGATLECGGPGWGN 376
Qy      355 DGRLLTPGJEDHYTLDMKLAMEBPPGPIILPIRVKDAEBAVALANKSPGLSSVFTPRD 414
      377 KGIVFPVFSNVTDERLAKKEIFGVQOIMFKSLDDVIDKRAANTFYGLSGVFTKDI 436
Db      415 OKAFDIANKLEVTGVTHINNKTRGPNFPPLGLKSGAGVQGIYSIEAMTVKSIVL 472
      437 DKALITISSALQACTVAV-NQYGVVSAQCPRGFRMGNGELGEYGHETHEVTKYTV 493
```

```
RESULT 17
US-09-602-777A-182
; Sequence 182, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroeder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
```

```
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 182
; LENGTH: 497
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-09-602-777A-182
```

Query Match 27.2%; Score 650; DB 4; Length 497;

Best Local Similarity 33.3%; Pred. No. 3e-52;

Matches 162; Conservative 86; Mismatches 201; Indels 38; Gaps 11;

```
Qy      8 YNNGEW--KSSVNOQIEILSPIDSSLSGFVPMTREREDHAMKAGREALPAMALTYEBA 65
      21 YINSGVWKAEGTQRNPVDPVAVQEWGVSYPEATASLDSAVGAARLALSMALTAERT 80
Qy      66 OYLHKAAADIIEERDEEIAIVLAKEISKAYNASVTEVTRADILRYAAEEGIRLSTADBG 125
      81 GYLKIKATIEBSRSEALALTNRENGSPISETRGAASNAAGIFRYA----TLAPLDOE 136
Db      126 GKMDASTGKKAIVIRQPIGIVLAIPYVYPNVLSSKTAAPALIGNVMPFPPTQSGVS 185
      137 DIRFPAGSAESIVDKDPGVICALIAPMNPINLVITKLAIPALLAGCTVYIIRPASPTPLS 196
Qy      186 GLVIAKAFPAEAGLPAGVFNTITGRGSEIDYIYEHEVEVNFINTGSTPVGORIGKLAG-- 243
      197 IRFIIETAEAAVGPAGVNVLLTGSQ-RFGDALVRHGVDKVAFPGSTPVGKKIAAACGEL 255
Qy      244 MRPIMLELGGKAGIVLADADLDNNAKOIVAGAYDYSQORCTAIRVLVEEVDLIAEK 303
      256 IAPVTLLELGGKSAIILPDADMSVLTSLIRSCMNTGTCTYISRIIAPSSRYAEVQGT 315
Db      304 ISENVAKLSVGDPFDNATV-TPVIDNSADPIESLVDAKOKAK-----ELAEFKR 353
      316 VASTIAAGROGDPYDEETVFGPVASQYSTWMSYIDSAREGARVAVAGTSSISISEGL 375
Qy      354 RGRLLTPGLPDHYTLDMKLAMEBPPGPIILPIRVVD--AEENVAIANKSPGLSSVFT 410
      376 ESSEGIQTPVADVPDMKISREELTFGPVISTLKVDYDTNGVEBALANNTKFGGLGLVF 435
Qy      411 TRDFOKAFDIANKLEVTGVTHIN---NKTGRGPNDFPPLGLKSGAGVQGIYSIEAMT- 465
      436 GADDEGALAEVAVQVSGSVGINFPGSNHSA-----PFGRRHESGNGVE---YGIKGLSA 486
Qy      466 --NVKSI 470
      487 YLTYSKI 493
Db
```

RESULT 18

```
US-09-538-092-1139
; Sequence 1139, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1139
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P47895
US-09-538-092-1139
```

```
Query Match      27.2%; Score 650; DB 4; Length 512;
Best Local Similarity 34.0%; Pred. No. 3.2e-52;
Matches 164; Conservative 85; Mismatches 208; Indels 26; Gaps 9;
```

```
QY      8 YNNGWKSQSVN--QIEILSPIDSSLGVPVPMATREBVNDHAKAGREAL---PAMALIVY 62
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      35 FINNEMHESKSKKFKATCNPSTREQICEVEBGDKPVDKAVBAQVAFQSGPWRRLDAL 94
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      63 BRAQYLHAKRADIIEERKEEIAIVLAKEISK-----AYNASVTEVVRTADIRYAAEGIRL 118
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      95 SGRLLHLQADLVERDRATLALLETMDTGKPLHAFIDLEGCIPT--LRYFAGWADKI 151
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      119 STSADGGMKMDASTGHLKLAIVIRQPVGIATAPVYVNVNLGSKIPALIGNVVMFKP 178
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      152 Q-----GKTIPTDNNVCFTRHEPIGVCGAITPNNFPLMLVWKLAPALCCGNTMVLKP 205
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      179 PTQGSVSGIVLAKAFAGLPAAGVNTITGSGSEIGDIYVEHEVNFNFTGSTPVGQRI 238
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      206 AEQTPLTALYLSLILKEAGFPFGVNVIVPGFPTVGAALISHPQINKIAFTGSTEVGKLV 265
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      223 GKLA---GMRPIMLELGGKDAIVLADADLNNAAQIYAGADVSGQRTAIKRYLYVEE 295
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      266 KEAASRSNLIKRVLTLELGGKNCPCIVCADADLDLAVECAHQGVFFNOGOCCTAASRVFVEEQ 325
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      296 VADELAEKISENVAKLSVGDPPDNATVT--PVIDNSADFIESLIVVDARQKAK--ELNE 351
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      326 VYSEFVRSEVYAKKRPVGDPPDVKTGEGPQIDQKQFDKILIELISGKKEGAKLECGGSA 385
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      352 FRDGRLLTPGLFDHVTITMDKLAMEBPPGILPIIRVKDAEBAVALANKSPGLQSSVPT 411
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      386 MEDKGLFKPTVFSEVTNMRILAKEEIFGPVQPIILFKFSIEEVIKRA NSTDYGLTAAVPT 445
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      412 RDPQAFADLANLLEVTGTHINNKTRGPDNPPFLGLKSGAGVQGIIRYSIEAMTVKSIY 471
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      446 KNLDRALKLASALBSGTWINCYNALYA-QAPFGGFKMSGNGRELGEYALAEYTEVKYVT 504
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      472 LDM 474
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      505 IKL 507
```

```
RESULT 19
US-09-961-403-12
; Sequence 12, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMP, HOLGER
```

```
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-12
```

```
Query Match      27.2%; Score 650; DB 4; Length 512;
Best Local Similarity 34.0%; Pred. No. 3.2e-52;
Matches 164; Conservative 85; Mismatches 208; Indels 26; Gaps 9;
```

```
QY      8 YNNGWKSQSVN--QIEILSPIDSSLGVPVPMATREBVNDHAKAGREAL---PAMALIVY 62
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      35 FINNEMHESKSKKFKATCNPSTREQICEVEBGDKPVDKAVBAQVAFQSGPWRRLDAL 94
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      63 BRAQYLHAKRADIIEERKEEIAIVLAKEISK-----AYNASVTEVVRTADIRYAAEGIRL 118
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      95 SGRLLHLQADLVERDRATLALLETMDTGKPLHAFIDLEGCIPT--LRYFAGWADKI 151
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      119 STSADGGMKMDASTGHLKLAIVIRQPVGIATAPVYVNVNLGSKIPALIGNVVMFKP 178
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      152 Q-----GKTIPTDNNVCFTRHEPIGVCGAITPNNFPLMLVWKLAPALCCGNTMVLKP 205
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      179 PTQGSVSGIVLAKAFAGLPAAGVNTITGSGSEIGDIYVEHEVNFNFTGSTPVGQRI 238
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      206 AEQTPLTALYLSLILKEAGFPFGVNVIVPGFPTVGAALISHPQINKIAFTGSTEVGKLV 265
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      223 GKLA---GMRPIMLELGGKDAIVLADADLNNAAQIYAGADVSGQRTAIKRYLYVEE 295
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      266 KEAASRSNLIKRVLTLELGGKNCPCIVCADADLDLAVECAHQGVFFNOGOCCTAASRVFVEEQ 325
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      296 VADELAEKISENVAKLSVGDPPDNATVT--PVIDNSADFIESLIVVDARQKAK--ELNE 351
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      326 VYSEFVRSEVYAKKRPVGDPPDVKTGEGPQIDQKQFDKILIELISGKKEGAKLECGGSA 385
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      352 FRDGRLLTPGLFDHVTITMDKLAMEBPPGILPIIRVKDAEBAVALANKSPGLQSSVPT 411
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      386 MEDKGLFKPTVFSEVTNMRILAKEEIFGPVQPIILFKFSIEEVIKRA NSTDYGLTAAVPT 445
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      412 RDPQAFADLANLLEVTGTHINNKTRGPDNPPFLGLKSGAGVQGIIRYSIEAMTVKSIY 471
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      446 KNLDRALKLASALBSGTWINCYNALYA-QAPFGGFKMSGNGRELGEYALAEYTEVKYVT 504
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      472 LDM 474
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      505 IKL 507
```

```
RESULT 20
US-09-221-294-2
; Sequence 2, Application US/09221294
; Patent No. 6268138
; GENERAL INFORMATION:
; APPLICANT: Riccardo Dalla-Favera and
; APPLICANT: Alessandro Masiimo Gianni
; TITLE OF INVENTION: A Retroviral Vector Capable of Transducing the
; TITLE OF INVENTION: Aldehyde Dehydrogenase-1 Gene and Uses of Said
; TITLE OF INVENTION: Vector
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
```



```

; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM 330 466 DX2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,294
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 42990-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-221-294-2

```

Query Match 27.0%; Score 645; DB 3; Length 521;

Best Local Similarity 34.1%; Pred. No. 9,6e-52;

Matches 164; Conservative 85; Mismatches 210; Indels 22; Gaps 10;

```

QY 8 YNNGEKSSVN-QIILSPIDSSIGFVPAMTRREYDHAMKAGREAL-PAMALTY 62
DB 40 FINEWHDSGKGFVFNPAETEEELCOVEESGDEVDKAVKARAFQIGSPWRMTDS 99
QY 63 ERAQYHKAADIIBRDKEIATVLAKEI--SKAY-NASVETVRTADLIRYAABEGIRL 118
DB 100 ERRLILYKLDLIERR-LIATWESWESMNGKLYSNAVINDLAGCITLIRYCAWADKI 158
QY 119 STSAGEGKMDASTGHKLAIVIRQPVGIVLAIAFPYVPVNLGSKIAPALIGNVVMFPR 178
DB 159 QGQ---GRITPIDGNFFYTRREPIGVCGQIIPWPLVLMWKIGPLSCGNTVVVPR 214
QY 179 PTQGSVSGVLAKAPAEGLPAGVFNITIGRGEIIGDYIYHEEVNFINTGSTPVQGR 238
DB 215 AEDTPTLTAHVASLIEAGFPFGVNIIVGYPGTAAGALISHMDIDKVAFTGSTEVGKI 274
QY 239 GKLAG---MRPIMLEIGSKDAGIVLADADLDNAKQIVAGAYVSGORCTAIKRVLYVE 295
DB 275 KEAAGSNLRKRVTLLEGSKSPCTVLADADDNAVEPAHHGVFHHQGCCIASRIVEES 334
QY 296 VADELAEKISENVAKLSVGDPE-DNATVTPVIDNSADFTESLVVDAROKGAK---ELNE 351
DB 335 IYDEPVRREVERKAKYILGNPLTPGTQGHQIDKEQYDKLIDIEGKKEGAKLEGGGP 394
QY 352 FKDGGLLTPGLFDHVTLDMLKAMEEPFGFPIIRVKAEEVAVALANKSDFLGSSVFT 411
DB 395 WGNKGYFVQPTVFSNVTDEKRIAKEIFGVPQIMKFSIDVDYIKRANNTFYGLSAGVFT 454
QY 412 RDPQKAFDIANKLEVGTVHNNKTGPGDPNFPPLGLKSGAGVQGIYRISIEANTNKSIV 471
DB 455 KQIDKAITTSSALQAGTVWV-NCYGVVSAQCPFGGRMSGNGRELGEYGFHEYTEVKT 513
QY 472 L 472
DB 514 V 514

```

RESULT 21
US-09-489-039A-12722
; Sequence 12722, Application US/09489039A
; Patent No. 6610835
; GENERAL INFORMATION:

```

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12722
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12722

```

Query Match 26.8%; Score 642.5; DB 4; Length 507;

Best Local Similarity 32.9%; Pred. No. 1,6e-51;

Matches 162; Conservative 99; Mismatches 196; Indels 35; Gaps 13;

```

QY 4 EYQNYNGEKSSVN-QIILSPIDSSIGFVPAMTRREYDHAMKAGREALPAMALTY 61
DB 23 EQLYINGGYTSATSGRTFETINPATGEVLAIVQAGREDVDRAVESAQGGKIVAMAMTA 82
QY 62 YERAOYLHKAADIIBRDKEIATVLAKEISKAYN-ASVTEVVRTADLIRYAABEGIRLST 120
DB 83 MERSRLRAVDLRLQRNDELARLETLDGKPLSETAAVDIVTGTDLLEYA--GL---I 137
QY 121 SADEGKM---DASTGHKLAIVIRQPVGIVLAIAFPYVPVNLGSKIAPALIGNVVMFK 177
DB 138 PALBSGQIRLARS---FVYTRREPLGVAGIGAMNVPYIQALMKSAVALAAGNMIK 193
QY 178 PTQGSVSGVLAKAPAEGLPAGVFNITIGRGEIIGDYIYHEEVNFINTGSTPVQGR 237
DB 194 PSEVTPPLTKLARIYSBAGLPDGVFNVLPGIAGETGYLTGHPIDAKISFTGVASGK 253
QY 238 I---GKLAGMRPIMLEIGSKDAGIVLADADLDNAKQIVAGAYVSGORCTAIKRVLYVE 294
DB 254 VANSBAASLKEVTELGGKSPLIADDDLDLADIAAMNPFSSGGVCTGTRFVPA 313
QY 295 EVADELAEKISENVAKLSVGDPE-DNATVTPVID---DNSADFTESLVVDAROKGAKEL 349
DB 314 KKAERHKLIERVARIRAGDLFADDTNGPLVSPPHRNVALRYES---GKREGARLL 369
QY 350 ---NEFKD---GRLLTPGLFDHVTLDMLKAMEEPFGFPIIRVKAEEVAVALANKSD 402
DB 370 CGEALKGDGPDNGMVAFTVFTDCSDEMTIVREEIFGVPVMSILSYADEAEVIRANATE 429
QY 403 FGLQSSVFTRDQKAFDIANKLEVGTVHNNKTGPGDPNFPPLGLKSGAGVQGIYRISYE 462
DB 430 YELAGVVTVPNLRHRLIHOLEAGICWINS-WGESPAEMPVGKIHSGIGRENGVMTQ 488
QY 463 AMTNVKSIVLDM 474
DB 489 STYQVKSIOVEN 500

```

RESULT 22
US-09-949-016-9207
; Sequence 9207, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO: 9207
: LENGTH: 532
: TYPE: PRT
: ORGANISM: Human
US-09-949-016-9207

```

Query Match	26.7%;	Score 639;	DB 4;	Length 532;
Best Local Similarity	33.7%;	Pred. No. 3.6e-51;		
Matches 162;	Conservative 81;	Mismatches 212;	Indels 26;	Gaps 9

```

OY      8 YNKGMSKSVN---QIEIISPIDDSSLSGFAPMAMRESEVDHAMKRGAL---PAMALITY 62
Db      55 PINNEMQNSSEGRVFPVYNPATGEQYCEVQEADKADIDKRAQARLAFISIGSWRRMDS 114
OY      63 ERAQYILHKAADIIEERKEBIATVLA---KEISKAYNASVTEVKTADLIRYAAEGIRL 118
Db      115 ERGRLLDKLADIYERDRRAVLATMESISNGGKPFQAFYVDLQGIKT---FRYYAGWADKI 171
OY      119 STRADEGGMADSTGHKLAVIRROPQVILATAIPVNYAPVNLSSKTAPLIGNNVMFKP 178
Db      172 H-----GWTIVDDGYFTFTTHEPIEGVCGQIIPMFPPLTMEFAMKTAIPALCCNTVIRK 225
OY      179 PTQGSVSGVLAKAFPAEAGLPAGVFNTITIGRGEISIDYIYEHBEVNFINEFTSGTPVGOIR 238
Db      226 AEQTPLSALYMGALIKKEAGFPFGVINILPGYGPITAGALASHIGIDIKIAFTSGTEVGKLI 285
OY      239 GKLAC---MRPIMLEIGGDAGVLADADLDNNAKQIVAGAYDYSQORCTAIRKVLVYEE 295
Db      286 GSAARSNIKRYTLEIGGKSPNITFEADDLIDVAVEHAQGVFPNQOCCTAGSRIIVEBS 345
OY      296 VADELAEKISENVAKLSVSDPEFDNAVVT--PVIDDNGADPFIESIYVVAROKGASELNEFK- 353
Db      346 IYEEFVRRVERAKRRVVGSPFDPTTEGGQIQDKQYNKILBELIGSVAGALTEGGKG 405
OY      354 --RDGRLLTPGLFDHYTLDMKLAMEEPPGPIILPIIRYKAAEENVALANNSDFGLQSSVFT 411
Db      406 LGRKPFLEIPVFSNVTDMRIAKEBIFGPVBEILPFTKMTDEVIERANNSDFGLVAAVFP 465
OY      412 RDPQARFADNKLLEVQTVHINKKTGGRPNPFLGLGKSGAGVQGYRYSIEANTVNSIY 471
Db      466 NDINKALITYSSAMQGTWVINCINMALNAOS--PGGFRMGSGNGEMBEFGLREYSVKTYV 524
OY      472 L 472
Db      525 V 525

```

```

RESULT 23
US-09-489-039A-7569
; Sequence 7569, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7569
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7569

```

Query Match	26.5%	Score 635;	DB 4;	Length 498;
Best Local Similarity	33.1%;	Pred. No. 7.8e-51;		
Matches 161; Conservative	99;	Mismatches 197;	Indels 30;	Gaps 11

Oy	1	LTXYONXVONXGKSSVNO - IELISPIDDSLSGVPMTRBEEDHMKGRELP -	AA	56
Oy	19	LAIEIRLFINGEIICAAADNTTETIDPAAOQTLAONAGKADVERAVKARQOFNDGW		78
Oy	57	AALTVERAQYLHKAADIIERDKERIAVTLAKEISKAYNASV - TEVVRTDLIRYAAEG		115
Oy	79	SQSPQORKAILTRFADILMEAREHEILALLETDITGKPIRHSIRDDIPQAAAIRWAE -		136
Oy	116	IRLSTADE - GGRMDSTGHKLAIVRQPVGVLAAIPYNPVNLSSGKIARALIGNV		174
Oy	137	-----ALDKYGVAVTGSNELAMTVRETIQVIAAVWMPNPLLIACCKGPPALAAANSV		191
Oy	175	MFKPOTGSVSGVLAVAKAFEAEGPLAGVFNITITGKSEIIGUYIEHBEVNFINTGSPV		234
Oy	192	ILMPSEKSPLLTRALRAGLAKAEGRLDDGVLVNVSQGHGABQELALHPDVEITPTGSTR		251
Oy	235	GQRIKGLAG---MRPMLTELGSKQAGIYLADA - DLDNAAKIUVAGADVDSQORCTAIKRV		299
Oy	252	GKOLLNDAGSNKRRWLEBAGSKSNIVAFDCEPDLQCVNRATAGIIFPNQOVCITATRL		311
Oy	291	LJVEBEVADLEAKESISENVAKLSVGPDP - NATVTFPIVDINSADEIESLVAVDAROKAGEL		349
Oy	312	ILBESITADEFLANILKAEQHMGQGNRPDRDPTMGMLINDTHADNVHSI -		RGESQ 366
Oy	350	NEFKRDR-----LITPGLPDTLDMKLANEERPGIETIRYKAEAEVAILANKSD		403
Oy	367	STPLGGRKNPMAAGPTIFVDVDPASTLSHEELFGVILVVTFRKEEELKLANDSDY		426
Oy	404	GLQSVFTRDPOKAPVIANKLEVGTVHINNKGRRDNPPLGLGKSGAGVQOQIRYSIEA		463
Oy	427	GLGAAWTRDLASAHMSRRLKAGSVFVNNVND - GDMVFPFGVKQSGNSGNSIAHLEK		485
Oy	464	MTNVKSI 470		
Oy	486	FTIEIKTI 492		

```

RESULT 24
US-09-543-681A-7151
; Sequence 7151, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7151
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7151

```

Query Match	26.4%	Score 632.5;	DB 4;	Length 505;
Best Local Similarity	32.7%;	Pred. No. 1.4e-50;		
Matches 159;	Conservative 97;	Mismatches 209;	Indels 21;	Gaps 10

```

QY 3 KEI QNVYNGHMKSSVNVQ--IELLSPLIDSSSLGFPNPMREEVDPMKAGRALPAMALTT 60
Db 27 ESKYMLIGGEWWSSTQQTTKTYSPATGBELLSEYPSGTAEDVDLA VKAARAKAFPHMKRTS 86
QY 61 VYERAOYLHKAAADIIIRDKEEIAIYLAKEISKANYAST--EVRTADLIRAAABEGIRLS 119
Db 87 PAQRQALLIKLIDRLREOEERFAMWTESLDNGKPLNESNIIDLPASIDHFYFA--GV-IR 143
QY 120 TSADEGSKNDASGHKHLAVIRQPIGYIALAPVYVYVNLSSGKLAAPLIGNVYMKRPP 179
Db 144 SHHDESAYVDENT---LSLVIREPIGVGOIIPNMFPLPAMAKIAPAIAGDGVVNPFA 200

```



```
QY 119 STSADBGKMDASTGHTLAV-----IRQPGIVLAIAPYNPVNLSSGIAPALIG 171
DB 138 -----GLVPALIEGEQIPLETSFYVTRREPLGVAGIAMNYPVOIALMKSAAPLAA 190
QY 172 NVWFKPPTGSGVGLAKAPABAGLPAGVFNITITGRSGEIGDYIYEHEENVFINPTGS 231
DB 191 NAMIIFRSEVTPPLALKLAIITYEAGVPDGVFNVLISGREVGOMLTHEPLIKISTFG 250
QY 232 TPVGQRI--GKLGMRPIMLEIGKQDAGIVLADADLDNAKOI VAGAYDYSQORCAIK 288
DB 251 TSGKKVMASASSSLKEVITMELGSKSPLIIFPDADLDRADIAVAMNPFSSQGVCTNGT 310
QY 289 RVLVEEVADELAEKISENVAKISVGDPPD-NATVTPVIDNSADFIESI LVDAKQKAK 347
DB 311 RVPIHRSQARFPAKIVLERVQRI RLGDPODENTNFGPLVSPHMESVLGYESGKQAKAR 370
QY 348 ELNIEFKR-----DGLLTPGLFDHYTLDMKLAMEEPFGPIPIIRVKDAEBAVALANK 400
DB 371 LLOGGERVTDGANGNAYVAPVFTDSCSDMTIVREEIFGPVMSILYDDEDEALIRAND 430
QY 401 SDFGLOSVEFTRDFOKAFDIANKLEVGTVHINNKTRGPDNFPPLGKSGAGVQIGIRYS 460
DB 431 TBYGLAAGVVTQDLARAHRAIHRLBAGICWINT-WGESPAEMPVGQYKQSGVGRENGLLT 489
QY 461 IEAMTVKSIIVLDM 474
DB 490 LAHYTRIKSVQVEL 503

RESULT 30
US-09-134-001C-4541
; Sequence 4541, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stramm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4541
; LENGTH: 510
; TYPE: PR1
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4541

Query Match 26.1%; Score 624; DB 3; Length 510;
Best Local Similarity 31.8%; Pred. No. 8.7e-50;
Matches 155; Conservative 101; Mismatches 209; Indels 22; Gaps 10;
```

```
DB 256 VAOAGAERIVPTTLIELGSGSANIIFPDANLEQVIEBVOGILFNGEVSACSARLLVQSS 315
QY 296 VADELAEKISENVAKISVGDPPDNAT-VTPVIDNSADFIESI LVDAKQ-----KGAK 347
DB 316 IYVELLPKLKEAFENIKYVDDPFDEDTKMSAQSGPEQDLKIESYIKIAEEDKANILITGGH 375
QY 348 ELNIEFKRD-GRLLTPGLFPHVTLDMKLAMEEPFGPIPIIRVYDAEBAVALANKSPGLQ 406
DB 376 RTDNGLDKGYFEEFPIIIEINDKQLAQEELFGPVVVEKEDDECEALEIANDSESGULA 435
QY 407 SSVFTRDFOKAFDIANKLEVGTVHINNKTRGPDNFPPLGKSGAGVQIGIRYSIAMTN 466
DB 436 GGIFTTDHRLALNVAAMKTRGIWINT-YNOIPAGAPFGQYKSKGIGREYVKDALIKYQO 494
QY 467 VKSIYLD 473
DB 495 VKNIFID 501

RESULT 31
US-09-328-352-6753
; Sequence 6753, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6753
; LENGTH: 493
; TYPE: PR1
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6753

Query Match 26.0%; Score 623; DB 4; Length 493;
Best Local Similarity 32.8%; Pred. No. 1e-49;
Matches 188; Conservative 94; Mismatches 214; Indels 16; Gaps 6;
```


Query Match	26.0%	Score 621;	DB 3;	Length 508;
Best Local Similarity	34.0%;	Pred. No. 1.7e-49;		
Matches	156;	Conservative	80;	Mismatches 205; Indels 18; Gaps 7.
Qy	23	LSPIIDSSIGFPAATREVEVDAMKAGEALPAMALVTVEBAQVYLHKAAIDIERDKKEI	82	
Db	51	INPADGSHLASAEKATADVARAVAAKAAATWQRMKPAQTRIMFRYALIEBKTEL	110	
Qy	83	ATVLAKEISKAYNASV-TEVVRTADLIRYAAEEGIRLSTSADEGKMDASTGHKLAVIR	141	
Db	111	AQLOGRDMGKPIRESLIGIDLPIMIETLEFFA----GLVTKIE--GRTPAPGRFLNTLIR	164	
Qy	142	OPVGIVLAIAIPNTYVNNISGSKIAPALIGGVNMKRPPTQGSVSLVAKAPAEKLPAG	201	
Db	165	EPIGVGAITPNNFPFAVOAWMKIAPALMGNAIVLKPQALPLVVALGELIAELGALPPG	224	
Qy	202	VFNITITGSGEIGDYIVHEEBVNFINFPGSPVGGORICKLAGMRPI--MLEIGGADAGIV	259	
Db	225	LNVNLPDGGSVAGNALVQHPSPGKATFTGSTEVGQIGGMADRLITASLEIGGGSALVA	284	
Qy	260	LADADLDNAAKQIVAGAYDYGQRCETAIKRVLVEVADELAEKISENVAKLSVGDPPDN	319	
Db	285	FQDSSPKXAAVAVPQAMTSGNGEITTAISRLLIVERPIYDEVVELVQARVEARVDPDIP	344	
Qy	330	AT-VTPVIDNSADPIESLVDARQKAKEIN-----EFKRDGRLLTPGLFDHYITDM	371	
Db	345	DTEIGPLISAEQRESVHSVVGSTEGEATLLISGDPSPFGAPEQGFYRPTLFGSVTADM	404	
Qy	372	KLAMEEPPGPIIPITRVDAEAAVLAANKSDPGLQSSVPTPDFOQAFLIANKLEVGTVHI	431	
Db	405	RIAREEIRGPVLSVLPFGEEBEAITLADTVAFGLAAGVFTTRVGRALRPAQTLIDAGNWI	464	
Qy	432	NNKTGRGPDNPFLGLKSGAGVGQIGIRYSIEAMTVKSI	470	
Db	465	NSMGVLANPAS-PYRGFGQSGYSDLGQAAIESFTKEKSI	502	
RESULT 35				
US-09-538-092-851				
Sequence 851, Application US/09538092				
Patent No. 6753314				
GENERAL INFORMATION:				
APPLICANT: Glot, Iolc				
APPLICANT: Manfield, Traci A.				
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same				
FILE REFERENCE: 15966-542				
CURRENT APPLICATION NUMBER: US/09/538,092				
CURRENT FILING DATE: 2000-03-29				
PRIOR APPLICATION NUMBER: 60/127,352				
PRIOR FILING DATE: 1999-04-01				
PRIOR APPLICATION NUMBER: 60/178,965				
PRIOR FILING DATE: 2000-02-01				
NUMBER OF SEQ ID NOS: 1387				
SOFTWARE: CuratPatSeqFormatter Version 0.9				
SEQ ID NO 851				
LENGTH: 517				
TYPE: PRT				
ORGANISM: Homo sapiens				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (0)...(0)				
OTHER INFORMATION: Polypeptide Accession Number P05091				
US-09-538-092-851				
Query Match	26.0%;	Score 621;	DB 4;	Length 517;
Best Local Similarity	33.7%;	Pred. No. 1.7e-49;		
Matches	162;	Conservative	85;	Mismatches 212; Indels 22; Gaps 10.
Qy	6	ONYVNGEKSSVNO--IELISPIDSSIGFVPMATREVDAMKKGREAL--PAMMALT	60	
Db	38	QIFINNEHMDASRKTFPPVNSSTGEVYICQVAEGDEVDKAVKARAAFPQSGSPRRMD	97	
Qy	61	VERAQVYLHKAAIDIERDKKEIATVLAKEISKAYNAS-VTEVVRTADLIRYAAEEGIRLS	119	

[illegible]

```

RESULT 36
US-09-949-016-10653
/ Sequence 10653, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO011307
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 10653
/ LENGTH: 520
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-10653

Query Match          26.0%; Score 621; DB 4; Length 520;
Best Local Similarity 33.7%; Pred. No. 1.7e-49;
Matches 162; Conservative 85; Mismatches 212; Indels 22; Gaps 10;

        6 ONVYNGEKKSVNQ--RIELSPIDDSGFWPDMTRREEDHMKAGREAL---PAAALT 60
        |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 41 QIFINNEHNDVSRKTFPTVPNPGSEYVLCQVAEGKEDVDKAKARAAAFOLGSEPRRRD 100
        |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

        61 YVERAQYLAKADIIERDKKEIATVLAKEISKAYNAS--VTEVVRTADILRYAAEGIRLS 119
        |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 101 ASHGRLINLRADILIERORTYLAALLETLDNGKPRYITSLVDVDMVLKCRYYA----- 153

        120 TSADE-GSKMDASTGCHKLAVIRRPQVGIIVLAIAPNYPVNUNSGSKTAPALLIGNTVMFKP 178
        |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
        154 GWAPKRYHKETIPIDGDFPSYTRHEVGVCGQIIPNNFPLMQAMWKLGPMLATGNTVMVKV 213

```

QY 179 PTOGVSGLVLAKAFAGLPGAVNTITGRGSEIGDYIHEHEVNFINTGSTPVGQRI 238
DB 214 AQTPLTALYVNLIKKAGFPFGVNIIVPGFPTAGALASHEDVDKVAFTSGTEIGRVI 273
QY 239 GKLAC---MRPIMLEIGSGDAGIVLADLDLNAKQIVAGAVDYSQRCRTAIKRVLYVE 295
DB 274 QVAAGSSNKRVTLEIGGSGSPNIMSDMDMAVEQHPALFENOQCCASRTVQED 333
QY 296 VADELAKISENVAKLSVGDPPDNATVT--PVIDDNGADFTESLVVAROKAKEL--NEF 352
DB 334 IYDEFERSVARAKSVVGNPDSKTEQGPQVDETQFKILGINTGQEGAKLGGGI 393
QY 353 KED-GELLTPGLFDHVTLLDMKLAMEBPPGPIILIRVKDAEBAVALANKSDGLOSVFT 411
DB 394 AADRGGFIQPTVEGVDQDGMTAKEEIFFGVMOILFKTIEVVGRRANNSTYGLAAVFT 453
QY 412 RPOKAFDIANKLEVTTTHIINKTGPPNPFPLGKSGAGYQGRISYEAMTNYKSIV 471
DB 454 KDLDRANYSQALQAGTVMWNCYDVGAQS--PFGYKMSGSGRELGVEYLOAYTEVKVT 512
QY 472 L 472
DB 513 V 513

RESULT 37
US-09-328-352-6687
; Sequence 6687, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6687
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6687

Query Match
Best Local Similarity 31.3%; Pred. No. 3,1e-49;
Matches 149; Conservative 97; Mismatches 206; Indels 24; Gaps 7;

QY 6 QNVNGEMKSSVNO--TELSPIDSSLGFPVPMTRREVDHANKAGREALPMAALTVYE 63
DB 29 QAFIAGKWDADHQOTSILNPATLEIGTVPWNGRAEAEALAEAKEMPLMKNTAKD 88
QY 64 RAQYLHKADIIERDKKEIATVLAKESKAYNASVTEVKTADLIRYABEGIRLSTSD 123
DB 89 RSILKKPDLIISNDELAFILTSBQKPLAEKELIYASFIEMFAEAKRYVGDIT 148
QY 124 EGGKMDASTGSHKLAIVRRQVGIIVLAPYNYVNSGSKIAPALIGNVVMPKPTQGS 183
DB 149 PEPYPPAR-----IYVKKPIGVAAITTPNPFPAAMITRKVAPALAAGPCIYKAPETP 203
QY 184 VSGVLAKAFAGLPGAVNTITGRGSEIGDYIHEHEVNFINTGSTPVGQRIKLAG 243
DB 204 FALALVLDLAVQAGVPAISFVSITGDVHIGDAIFESDVRKFTFGSTPVGMLERSA 263
QY 244 --MRPIMLEIGSGDAGIVLADLDLNAKQIVAGAVDYSQRCRTAIKRVLYVEVADLA 301
DB 264 KTLKRVSLBGGAPPIVFDADLDLAIREGALLAKRRNAGQTCVCRFLVQAGIYEKFI 323
QY 302 EKISENVAKLSVGDPPD-NATVTPVIDNSADFIESLVVAROKAKEL--NEFRDOR 357
DB 324 AALSQKIQNFNIGNGLEAGHDIGPLINAAVKKVEAHIDDLKNGRLVGGKKAAGEL 383
QY 358 LITPGLFDHVTLLDMKLAMEBPPGPIILIRVKDAEBAVALANKSDGLOSVFTRFOKA 417

DB 384 FPEPILLIADVADMDVATQETFGPLAAVKEETEQOAVEMANNATEFGAAAYCYTKDLGRA 443
QY 418 PDIAKLEGTGTHIN-----NKTGRGPNPFLGKSGAGYQGRISYEAMTNYK 468
DB 444 WMSQQLRYGVNGVINKGLISNEVA-----PFGGIKQSGLREGSKYGIEDYLEIK 493

RESULT 38
US-09-949-016-9559
; Sequence 9559, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9559
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9559

Query Match
Best Local Similarity 34.4%; Pred. No. 3,5e-49;
Matches 151; Conservative 79; Mismatches 184; Indels 25; Gaps 8;

QY 8 YVNGEMKSSVN--QIEILSPIDSSLGFPVPMTRREVDHANKAGREAL--PMAALTVY 62
DB 2 FINNEMHSSKSGKPKATCNPSRTEQICEVEBGDKPDVKAVEAQAVALQRGSPWRRLDAL 61
QY 63 ERAQYLHKADIIERDKKEIATVLAKESK---AYNASVTEVKTADLIRYABEGIRL 118
DB 62 SGRGLHOLADLYERDRATLALETMDYKGFHLAFIDLRGCIPT--LRYFAGMADKI 118
QY 119 STSADGGMDASTGSHKLAIVRRQVGIIVLAPYNYVNSGSKIAPALIGNVVMPK 178
DB 119 Q-----GKTIPTDNNVCFTRHEPTIGVCAITTPNPFPLMLWMKLAPLCCGNTMVLKP 172
QY 179 PTOGVSGLVLAKAFAGLPGAVNTITGRGSEIGDYIHEHEVNFINTGSTPVGQRI 238
DB 173 AQTPLTALYVNLIKKAGFPFGVNIIVPGFPTAGALASHEDVDKVAFTSGTEVGKLV 232
QY 239 GKLAC---GMRPIMLEIGSGDAGIVLADLDLNAKQIVAGAVDYSQRCRTAIKRVLYVE 295
DB 233 KEAAGRSMIKRVTLLEIGGNPCIVCADLDLVAECAGVGFNOGOCCTAASRVFVEBQ 292
QY 296 VADELAKISENVAKLSVGDPPDNATVT--PVIDDNGADFTESLVVAROKAKEL--ELNE 351
DB 293 VTSSEFVRSVVEYAKRPVGDPPDVTTEQGPQDQKQFQKILIELISGKKEGAKLECGGSA 352
QY 352 FRDRLTLPGLFDHVTLLDMKLAMEBPPGPIILIRVKDAEBAVALANKSDGLOSVFT 411
DB 353 MEDKGLFIKPIYFSVTNNMARIKEIEIGPVPILKFSIEBIVIKRANSTDVGLTAAVFT 412
QY 412 RPOKAFDIANKLEVTVH 430
DB 413 KDLDRALKLASALRESGTVA 431

RESULT 39
US-09-155-183-4
; Sequence 4, Application US/09155183
; Patent No. 6323011


```

Db 41 INGEKLTFTDIPFNSVNPANTSQTLIAKYSKATODDIEKAFESANHAHYOSWKMSHKDRBEL 100
Qy 68 LHKAAADIIENDKEBIATVLAKEISKAYNASYTEVTRTDLIRYAAEBSIRLSTSADEGK 127
Db 101 LLRAVAIIIRRKKEBISAIMVYEGAPMEDAVGDAAEIGDPIEYARSMMEL--ADGKPV 157
Qy 128 MDASTGHKLAIVIRPOVGIUALAPYNP-VNLGSKTAPALLIGANVMFKPPTQGSVSG 186
Db 158 LDREGEHNRKRYFYK--PIGTGVITIPMNPPAIMAGTTLAP-VVAGNTVLKPAEDYVTLTA 214
Qy 187 LVLAKAPREAPLPAQVFTTITGRGSEIGDYIYEHEEUNVFINFPTGSPVQSIGKLAC--- 243
Db 215 YKMEIIEEAGLPQGVNPFVPGDPEIGDYIVDHDHTFVFTQSRATGRIYERSAVVQ 274
Qy 244 -----MRPIMLELGKADGIYVLADADLDNAKQIYAGAYVDYDSGORCTAIKRVLVEEYAD 298
Db 275 EGQGFPLKRIYEMGGKDAIIVDDNNVDTDLAAEATVTSFGFSGQCSCSRIYAHQDVHD 334
Qy 299 ELAEKISENVAKLISGDPFDNATVTPVIDDMSADFIESLVVDAROKAKELNEFKRD--G 356
Db 335 ELERKAQQLTQKLTIGTNEENTFMGPFVINQOFQDKIKYIIEIKKKEGLLETFGGGTDDSTG 394
Qy 357 RLITPPLGFLDHYLTDMKLAMEEPPFGPILP IIRVKOAEBAVAIAJANSDFGLQSSVFTDRFOK 416
Db 395 YFIEPTIPSGQASADRIMQEBIEFGVVGFIKVKDPEDAIEVANDPDTGLGAVITNREH 454
Qy 417 AFDIANKLEVGTVAHNNK-TGRGPDNPFPLGLKSGS 451
Db 455 WIKAWNEPDVGNLYINRGCTAAVVGYPHFGFGFKMSG 490

```

```

1  RESULT 42
2  / Sequence 2, Application US/08513841
3  / Patent No. 5753481
4  / GENERAL INFORMATION:
5  / APPLICANT: Niwa, Mineo
6  / APPLICANT: Saito, Yoshimasa
7  / APPLICANT: Ishii, Yoshinori
8  / APPLICANT: Yoshida, Masaru
9  / APPLICANT: Suzuki, Hiromi
10 / TITLE OF INVENTION: No. 5753481el L-sorbose Dehydrogenase and No. 5753481el L-sorbose
11 / TITLE OF INVENTION: Dehydrogenase Obtained from Gluconobacter oxydans T-100
12 / NUMBER OF SEQUENCES: 22
13 / CORRESPONDENCE ADDRESS:
14 / ADDRESSEE: Obolon, Spivak, McClelland, Maier & Neustadt, P.C.
15 / STREET: 1755 Jefferson Davis Highway, Suite 400
16 / CITY: Arlington
17 / STATE: Virginia
18 / COUNTRY: USA
19 / ZIP: 22202
20 /
21 / COMPUTER READABLE FORM:
22 / MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
23 / COMPUTER: IBM PC compatible
24 / OPERATING SYSTEM: PC-DOS/MS-DOS
25 / SOFTWARE: MS-DOS Editor
26 / CURRENT APPLICATION DATA:
27 / APPLICATION NUMBER: US/08/513,841
28 / FILING DATE: 01-NOV-1995
29 / CLASSIFICATION: 435
30 / PRIOR APPLICATION DATA:
31 / APPLICATION NUMBER: UK 9304700.9
32 / FILING DATE: 08-MAR-1993
33 / PRIOR APPLICATION DATA:
34 / APPLICATION NUMBER: JP 241851/1993
35 / FILING DATE: 28-SEP-1993
36 / ATTORNEY/AGENT INFORMATION:
37 / NAME: NORMAN F. OBLON
38 / REGISTRATION NUMBER: 24,618
39 / REFERENCE/DOCKET NUMBER: 18-909-0 PCT
40 / TELECOMMUNICATION INFORMATION:
41 / TELEPHONE: 703-413-3000
42 /

```

```

1 TELEFAX: 703-413-2220
2
3 TELEX: 248855 OPAT UR
4
5 INFORMATION FOR SEQ ID NO: 2:
6
7 SEQUENCE CHARACTERISTICS:
8
9 LENGTH: 497 amino acids
10
11 TYPE: amino acid
12
13 TOPOLOGY: linear
14
15 MOLECULE TYPE: peptide
16
17 ORIGINAL SOURCE:
18
19 ORGANISM: Gluconobacter oxydans
20
21 STRAIN: T-100
22
23 FEATURE:
24
25 NAME/KEY: mat peptide
26
27 LOCATION: 1..497
28
29 IDENTIFICATION METHOD: experimentally
30
31 US-08-513-841-2

```

Query Match	25.4%	Score 609	DB 1	Length 497
Best Local Similarity	30.8%	Pred. No. 2.1e-48		
Matches 149	Conservative 101	Mismatches 213	Indels 20	Gaps 8
QY	3	KEYQNVYVGEKSSVNOQIEILSPIDDSLGVPVPMTRTEVVDHAMKAGREALP--AWAALT	60	
DB	14	REFGFPIDEGRKAGKDFPDRSSPHADVPVTRIPKCTREDLDEAVALAARRAENSGMAGLA	73	
QY	61	VYERAOYLHKADADIIERDKIEIATVLAKESKAYNASVTEVVRTADLLRYAAEBGIRLST	120	
DB	74	AADRAAVLTKAAGLLRERRDDIAYWEVLENGKPIQAKGEIDHICIAFEMAAGAARMH--	132	
QY	121	SADGCKKADASTGHL-AVIRROPGVGLALAPFNYPNVLSGSKAPLLTIGNVVMFMRP	179	
DB	133	-----GDTFNNIGGGLFGMVLREPIGVGGLTTPMNFPMILCERAPFLIAGGCTILVPR	187	
QY	180	TQGSVGLVAKAFAPAGBAGLPAGVFNNTTGRSGEISGDYVBEHEVNFINTGSTPVGGRIG	239	
DB	188	EYTSATILLLEIILADAGLPKGVFENVVGTIGRTYGOAMTEHQDIDMLSFFTSITGVKSCI	247	
QY	240	KLA---GMRPMLIEIGKDGAGIVLADADLDNAAKOIVAGAYDVSGORCTALKRVLVVEE	296	
DB	248	HAADSNMKKLGLELGGKNPILVFADSNLLEDAAADVAGISFNNQCCVSSRLIVERSV	307	
QY	297	ADELALEKTSENVAKLSVGDPPDNAT-VTPVIDDNSADPISLVDAROKKAKEL-----N	350	
DB	308	AEEKERLVVPMOEKIRIVGDDPPDPTQIGAITTEAONKTIILYIAKGAKEGAKLCCGGGIV	367	
QY	351	EPRKDGRLTTPGLPDHVTLLDMKLAEMEPFGPILPIIRYKDAEBAVALANKSDPGLQSSVF	410	
DB	368	DFGK-GGQIIOPTLFTDYKPSMGIADELFGVVLASFPHDYDEALALANDVYGLAASVW	426	
QY	411	TRDFOKAFDINKLEVGTVHINNKTGRGPDNFPFLYGLSGAGVGGIRYSIEAMTNVVS	470	
DB	427	SKDIDKALAVRRYRABAGFWMTNTIMSGPER-PLGGFQSGMGREAGLYGVBEVYQISV	485	
QY	471	VLD 473		
DB	486	HIE 488		

RESULT 43
 US-08-696-834-2
 ; Sequence 2, Application US/08696834
 ; Patent No. 5834263
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Niwa, Mineo
 ;
 ; APPLICANT: Saito, Yoshimasa
 ;
 ; APPLICANT: Ishii, Yoshinori
 ;
 ; APPLICANT: Yoshida, Masaru
 ;
 ; APPLICANT: Hayashi, Hiromi
 ;
 ; TITLE OF INVENTION: Method for Producing 2-Keto-L-Gulonic Acid
 ;
 ; NUMBER OF SEQUENCES: 48
 ;
 ; CORRESPONDENCE ADDRESS:
 ;
 ; ADDRESSSEE: Obion, Spiwak, McClelland, Maier & Neustadt,
 ;
 ; ADDRESSSEE: P.C.
 ;


```

QY 3 KEYONYVNGEMKSVNQIIEILSPIDSSLGFPVPMATREEVDAHAKGREALP--AWAALT 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 14 REFGFFIDEMWAKGDFDRSSPAHDVPVTRIPRCTRELDDEAVAAARRAFENGSMAGLA 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 VYERQOYLHKADIIERDKERIATVLAKETSKAYNASTEVEVTRADLIRYAEBSIRLST 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 AADRAVALLKKAAGLIERRDDIAYWEVLENGKRIISQAKGEIDHCIACFEMAAGAARMH- 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 SADEGKMDASTGHKL-AVIRROPVGIYLAIPYNPVNLSSGKIAPALIGNVVMFKXP 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 -----GDTFNNIGELFGVNLREPIVVGILTFPMNPFMILCERAPFIILASGCTLVKPA 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 TQGSVGLVLAKAFAPAGLPAGVNTITGRGSEIGDYVEHEVNFINTGSTPVGORIG 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 EYTSATVTLLEIILADAGLPKGVFNVTGTGRVGOAMTEHODIMLSTGSTGVGKSCI 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 KLA---GMRPIMLEGGKAGIYVLADADLDNAKQIVAGADYSGORCTAIRVLVEEV 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 HAAADSNIKKLGLIEGGKNPIVVFADSNLEDAADAVAFISFNTGCCVSSSRLIVERSV 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 ADELAKISENVAKLSVDPFDNAT-VTPVIDNSADFIESIYVDAROKAKEL-----N 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 AKPERLVVPKMEKIRKVDPPDPETQIGAITTEAQNKTLLDYIAKKAEGAKLLCGGGIV 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 351 EFKRDGRLITPGLFDHVTLLDMKLAMEBPFGPIILPIIRVKADEBAVAIANKSDFGLOSVF 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 DFGK-GQYIQPTLFTDVKPSMGJARDEIFGPVLASFHPTVDEALAINADTYGGLAASTW 426
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 411 TRDPKADPIANKLEVGTVHINKTGRGPDNPFPLGLKSGAGVCGIRISTEAMTVKSI 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 427 SKDIDKALAVTRVRAGRFWNTIMSGPET-PLGGFKOSGWRGAGLVGEVYQIKSV 485
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 471 VLD 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 486 HIE 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 45
US-09-118-317-2
; Sequence 2, Application US/09118317
; Patent No. 6197562
; GENERAL INFORMATION:
; APPLICANT: Niwa, Mineo
; APPLICANT: Saito, Yoshimasa
; APPLICANT: Ishii, Yoshiro
; APPLICANT: Yoshida, Masaaru
; APPLICANT: Suzuki, Hiromi
; TITLE OF INVENTION: No. 6197562el L-sorbose Dehydrogenase and No. 6197562el
; TITLE OF INVENTION: L-sorbose Dehydrogenase Obtained from Gluconobacter
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Obion, Spiyak, McClelland, Maier & Neustadt, P.C.
; STREET: 1755 Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS-DOS Editor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/118,317
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513,841
; FILING DATE: 01-NOV-1995
; APPLICATION NUMBER: UK 9304700.9
; FILING DATE: 08-MAR-1993
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: JP 241851/1993
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: NORMAN F. OBLON
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-909-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 497 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Gluconobacter oxydans
; STRAIN: T-100
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..497
; IDENTIFICATION METHOD: experimentally
; US-09-118-317-2

```

```

Query Match      25.4%; Score 609; DB 3; Length 497;
Best Local Similarity 30.8%; Pred. No.2,1e-48;
Matches 149; Conservative 101; Mismatches 213; Indels 20; Gaps 8;

QY 3 KEYONYVNGEMKSVNQIIEILSPIDSSLGFPVPMATREEVDAHAKGREALP--AWAALT 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 14 REFGFFIDEMWAKGDFDRSSPAHDVPVTRIPRCTRELDDEAVAAARRAFENGSMAGLA 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 VYERQOYLHKADIIERDKERIATVLAKETSKAYNASTEVEVTRADLIRYAEBSIRLST 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 AADRAVALLKKAAGLIERRDDIAYWEVLENGKRIISQAKGEIDHCIACFEMAAGAARMH- 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 SADEGKMDASTGHKL-AVIRROPVGIYLAIPYNPVNLSSGKIAPALIGNVVMFKXP 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 -----GDTFNNIGELFGVNLREPIVVGILTFPMNPFMILCERAPFIILASGCTLVKPA 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 TQGSVGLVLAKAFAPAGLPAGVNTITGRGSEIGDYVEHEVNFINTGSTPVGORIG 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 EYTSATVTLLEIILADAGLPKGVFNVTGTGRVGOAMTEHODIMLSTGSTGVGKSCI 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 KLA---GMRPIMLEGGKAGIYVLADADLDNAKQIVAGADYSGORCTAIRVLVEEV 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 HAAADSNIKKLGLIEGGKNPIVVFADSNLEDAADAVAFISFNTGCCVSSSRLIVERSV 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 ADELAKISENVAKLSVDPFDNAT-VTPVIDNSADFIESIYVDAROKAKEL-----N 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 AKPERLVVPKMEKIRKVDPPDPETQIGAITTEAQNKTLLDYIAKKAEGAKLLCGGGIV 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 351 EFKRDGRLITPGLFDHVTLLDMKLAMEBPFGPIILPIIRVKADEBAVAIANKSDFGLOSVF 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 DFGK-GQYIQPTLFTDVKPSMGJARDEIFGPVLASFHPTVDEALAINADTYGGLAASTW 426
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 411 TRDPKADPIANKLEVGTVHINKTGRGPDNPFPLGLKSGAGVCGIRISTEAMTVKSI 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 427 SKDIDKALAVTRVRAGRFWNTIMSGPET-PLGGFKOSGWRGAGLVGEVYQIKSV 485
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 471 VLD 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 486 HIE 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 46
US-09-134-001C-4451
; Sequence 4451, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

```


Db 337 QRIKIDPDLIEDRMRPDLIRPHLERVLGFPVKAAGKAVLCCGDIYVEDEPKLKDGY 396
QY 359 LTFGLFDHVTLDMLAMEEPFGLPIIRVKDAEBAVALANKSDFGLQSSVFTRDFOKAF 418
Db 397 MRPCVLTNCDDMTCVKEEIFGVPVMSILSFTBAEVLERANDTTFGLAAGVTRDIOGRAH 456
QY 419 DIANKLEVGTVHNNKTRGPDNFPFLGLKSGAGVQGIKRSIEAMTNVKSIVLDM 474
Db 457 RVAAELQAGTCFINN-YNVSVELPFGYKSGSGFGENGKRVTEIYYSQLTVCVEM 511

RESULT 51
US-09-949-016-7505
Sequence 7505, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 7505
LENGTH: 553
TYPE: PRT
ORGANISM: Human
US-09-949-016-7505

Query Match 25.2%, Score 602, DB 4, Length 553;
Best Local Similarity 31.1%, Pred. No. 1.2e-47;
Matches 148; Conservative 99; Mismatches 197; Indels 32; Gaps 8;

QY 23 LSFIDSSL-GFPMATR-----EEVDHAKAGREALPAAALTVVERAQLYK 70
Db 79 VERADSGTRKAPEDPTGRVIATFTSGKEKVALVQNAKAKTISQKSMERCRILB 138
QY 71 AADIIRDBELATVLAKESKAYNASVTEVVTADLIRYAAEEGIRLSTADEGKMDA 130
Db 139 AARIIRDEBELATVMECINNKSIFEARLIDISWQCLEYYA-----GLAASMAEBHQ 192
QY 131 STGHKLAVIRROPGVIVLAIPNYPNVNSGSKIAPALIGANVMFKPTQSGVGLVIA 190
Db 193 LPPGSGFYRRREPLGVCVIGANNYFPFOIASWKSAPALACGNAMVFKPSFTVSAALLA 252
QY 191 KAPAEGLPAGVNTITGRGSEIGDYIVEHEEYNFINFGSTVGQRIGLA--GMRPLM 248
Db 253 EYISEGVPGLVNVVQG-GAATGQPLCGHPDAKVSFTGSGVPTGKIMEMSKGKIPVT 311
QY 249 LELGSGAGIVLADADLDNNAKOIVAGAYDSQRCCTAIRVLVVEEVADELAEKISENV 308
Db 312 LELGSGSPILIFSDCMANNAVKALMANPFTQGVCCNGRVRVQGEIIDKFTBEVYKQT 371
QY 309 AKLSVDDP-FDNAVTTPVIDDNSADTIESLVDAKQKAKEL-----NEFKDGRU 358
Db 372 QRIKIDPDLIEDRMRPDLIRPHLERVLGFPVKAAGKAVLCCGDIYVEDEPKLKDGY 431
QY 359 LTFGLFDHVTLDMLAMEEPFGLPIIRVKDAEBAVALANKSDFGLQSSVFTRDFOKAF 418
Db 432 MRPCVLTNCDDMTCVKEEIFGVPVMSILSFTBAEVLERANDTTFGLAAGVTRDIOGRAH 491
QY 419 DIANKLEVGTVHNNKTRGPDNFPFLGLKSGAGVQGIKRSIEAMTNVKSIVLDM 474
Db 492 RVAAELQAGTCFINN-YNVSVELPFGYKSGSGFGENGKRVTEIYYSQLTVCVEM 546

RESULT 52
US-09-710-279-1030
Sequence 1030, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1030
LENGTH: 496
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-09-710-279-1030

Query Match 25.1%, Score 601, DB 4, Length 496;
Best Local Similarity 32.5%, Pred. No. 1.2e-47;
Matches 152; Conservative 95; Mismatches 189; Indels 32; Gaps 10;

QY 6 QNYVNGEW--KSNVQIRILSPIDSSLGFPVAMTREEVDHAKAGREALP--AMALTV 61
Db 10 RQYIDGWEVSSNKTTRDILINPYNOETITVAEGTKEVERAILLAKRFEGEGMSLET 69
QY 62 YERAOYLHKAADIIRDBKEIATVLAKESKAYNASVTEVVTADLIRYAAEEGIRLSTS 121
Db 70 EYRGKKVRAVADKIKENREELAKLETLDTGKLTLESYADMDIHVFWFA--GL--AD 124
QY 122 ADEGKMDASTGHKLAVIRROPGVIVLAIPNYPNVNSGSKIAPALIGANVMFKPTQ 181
Db 125 KQGGELINSPIPNNAASKVKEPVGAVTQITPNVYPLQASWXTIAPALATGCSLVKPSBI 184
QY 182 GSVGLVLAKAPAEAGLPAGVNTITGRGSEIGDYIVEHEEYNFINFGSTVGQRIGRL 241
Db 185 TPLTITRVLEMEVGFPGKGTINVLVAGSEVDVMSGHEEVDVLSFTGISTGKIMQ 244
QY 242 AG--MRPIMLELGGAGIVLADADLDNNAKOIVAGAYDSQRCCTAIRVLVVEEVADE 299
Db 245 AANHVTDALELGGKQNPNIIFDDADFELAVDQALNGYFHAQVCSAGSRILVHNDIKQ 304
QY 300 LAEKISENVAKLSVDDPDMAT-VTPVIDDNSADTIESLVDAKQKAKELNEFRDGL 358
Db 305 FEKALIDRVSKIKLNGFPQDTEMGPVISTARDKIEGYMEVAKKOGATIALIGRRPERE 364
QY 359 -LTFGLFDHVTL-----DMKLAMEEPFGLPIIRVKDAEBAVALANKSDFGLQSSVFT 411
Db 365 DLQAGLFEFPYITTCDSMRIVQSEVSGPVTVGCFADDEALIRLANDSITGLAGAIT 424
QY 412 RDOKAFDIANKLEVGTVHNNKTRGPDNFPFL-----GLKSGAG 453
Db 425 KQIGRAQVRANKLKGITWIN-----DFHHYFAQAPWGGYKQSGIG 465

RESULT 53
US-09-134-001C-4383
Sequence 4383, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964


```

OY 203 FTTINGRGEIJDYVEHEENVEINFTNGSTPVGORJGSKLA--GMRJIMLELGKAGIYL 260
Db 205 FRVVG--GAATQGFLOQHDPVAKVSTGVSFPGMKIMESAKIKRYTLELGKSLIIF 263
OY 261 ADADLDNAAKQIVAGAYDYSGORCTAIKRVLVEEVADELAEKISENVAKLSVGD--FDN 319
Db 264 SDCDDMNNVKGALMNFLOGOVCCNGTGVFOKELDKFTBEVVVQTORIKIGDPLELD 323
OY 320 ATVPEVIDNSADFTESLIVDAROKGAKEL-----NEFRDORLLTPGLFDEVTLD 370
Db 324 TRMGELINPHERLVGFVKVAKEQAKVLCGGDIYVEEDPRLKDDIYRPFCLTNCRDD 383
OY 371 MCLAMEEPFPLIPIIRVDAEBAVAIANKSPGLQSSVFTTRDFQKAFDIANKLEVTYVA 430
Db 384 MTCVKEEIEFGPVMSTLSPTEBAVLERANDTTFGLAAGVFTTRDIOAHHRVVALAQGTCE 443
OY 431 INNKTRGRGDPNFPFLGKSGAGVGCIRISIEFMTNFKSIVLDM 474
Db 444 INN-YNVSFVELPFGGKYKSGFGREGRVLTIEYSQKTKVCVM 486

```

```

RESULT 56
US-09-949-016-5979
; Sequence 5979; Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5979
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-5979

```

```

Query Match      25.0%; Score 599; DB 4; Length 493;
Best Local Similarity 31.2%; Pred. No. 1,8e+47;
Match 145; Conservative 98; Mismatches 191; Indels 30; Gaps 7

Qy 33 FVPAATMR-----EEVDHAMKAGREALPMAALTYERAQYLAHKADIIRDEKEI 82
   ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 31 PEPAAGRVATATTCGSEKEVINLAVQAKAAFKMSQSGMERKIRILLEAARITREDEI 90
   ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Qy 83 ATVLAKESKAYNASVTEVVRTADLIRYAABEGIRLSTSBEGKNDASTGHTLAVIRRO 142
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 91 ATMECINNGKSIPEARLDIDISQCLEYYA-----GLASMGVGEHILQJPGSGFYTRRE 144
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Qy 143 PVGIYVLAIAPIVNYVPVNLSSKTAIPALIGNVVMFKPPTQSGVGLYLAKAFAPAGLPAGV 2022
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 145 PIGVCCVGIAGAMVYFPQIAQMSKSPALACNANAFKSPPTPVSAIIIAEYISEAGVPPGL 204
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Qy 203 FMTITGRGSEIDDIYIEHEENVNFINTGSPVGOBRIGKLA--GMRPIMLEGGKDGIVL 260
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 205 FNVVGO--GAATQGFICQHPDVAKVSTGVSYPFGMKIMEMSAKIKVPTTLELGGKSPILIF 263
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Qy 261 ADADLDNAAKOIVAGAYDYSGQRCTAIKRYLVVEEVADEBLAEKISENVATLSVGD--FDN 319
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 264 SDCDMNNAVKGALIMANFLTQGGVCCNGTRFVQKEILDKTEEEVAVQOTQRIKIGDFLLIED 323
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Qy 320 ATVTPEVIDNSADFIESLVVADQKAKEL-----NEFRDGRLLTPELFDHVTILD 370
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

Db 32 TRMGPIINPHERVLVGFVKAKEGQAKYLGGSDIYVEDPKLDQGYMPCVLTNRDD 3633

Qy 371 MKLAMEEPGPILIPITIRVDAEAAVIAANKSDPGIQQSVFRRPQKAFDIAKLEVEGVH 430

Db 384 MTCVKEIRIGPVMSILSPDTAEVLERANDOTTFGLAAVFPRDQGRHRYVAELQACTF 4433

Qy 431 INNKTRGPGDNPFPGLKQSGAGVOCIRISIAEMTNVKSIVLDM 474

Db 444 INN-YVNSPELLPFGGYYKSGGFRNGVLTLEYSSQLTVCVEM 486

```

RESULT 57
US-09-328-352-6118
; Sequence 6118, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID-NO 6118
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6118

```

```

Query Match 25.0%; Score 598; DB 4; Length 490;
Best Local Similarity 33.1%; Pred. No. 2,36-47;
Matches 151; Conservative 95; Mismatches 210; Indels 14; Gaps 6;

QY      12  EWKSSVNOEITLSPIDSSIGFVPMATREEDHAMKAGREALPMWALTYYERAOYLHKA 71
      13  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      26  EAPASNYIIEVNDATGETTLLAVKTYPRAGVEAALINSSAKQAAMKQOTALAPRADVLIAW 85
      27  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      72  ADIIEBDKEEIAIYLAKEISKAYNASVTEVVRTADLIRYAAEEGIRLSTADEGGKMDAS 131
      73  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      86  YNLMLEHKENLAQILTBQCKPLAEAGEIGIYAASFIRWPEQARRI-----DGEVLTP 140
      87  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      132  TGHKLAVIRRPQVIGVIALAIEYNYPNVLSSKIALPALIGNVNMFKEPTQSGVSGLVLAK 191
      133  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      141  LPNQRLLVIKQAIQVTAITPMNPAMITKAPALIAAGCSMLVKEPEQPTLYALIEV 200
      142  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      192  AFAAGAGPAGFENITIGRSFEGIYIIEHEHYVNIINFTGSPVQORIGKLAG--MRPIML 249
      193  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      201  LAIDAGLRADVLINISGDPAVEVGKTLCSBDIVKLSFSGISQVERILMQOQAPITKLSL 260
      202  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      250  ELGGKDGIVLADADLDNMAKQOIVAGAYDYSGORCTAIKRVLYVEEVADELAEKISENVA 309
      251  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      261  ELGGNAPVYVDDANLBEQAVGIMASKYRNSGQCTVCANRIYVDGIIYDALADRLVEAS 320
      262  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      310  KLSIVGD--PFDNATYTPVIYDINSADPIESLSVDAQOKAKELNEKRDP---GRLLTPGLFD 365
      311  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      321  KLVKGDGRQESSTGPIIDEDALAKVOSHADATEKATYRIGGKRKALGTFEFPVLT 380
      322  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      366  HTVLDMLKLAEEPPGPIPIIRVDAEAEVAALNAKSDFGQSSEVTFEDFOKAPDIANKLE 425
      367  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      381  GVTQDMKYSKEETGPIAPLFRFRFTBEDEAVAMANDTEFGLAALYFTQSTANQWVGALAE 440
      382  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      426  VGYTHINNKTKGR--GPDNPFPIGLKSGAGVQGIKYSIEAMTNVYSIYLDM 474
      427  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      441  YGMVGIN--TGAISENVAAPFGGVKQSGLGREGSKFGLIEYEMKYLCLVDL 488
      442  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 58
US-09-717-926-2
; Sequence 2, Application US/09717926
; Patent No. 659657
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Cook, William J.

```

```
/ TITLE OF INVENTION: 32140, A No. 6569657el Human Aldehyde
/ TYPE: PRN
/ FILE REFERENCE: 35800/205243
/ CURRENT APPLICATION NUMBER: US/09/717,926
/ PRIOR FILING DATE: 2000-11-21
/ PRIOR APPLICATION NUMBER: US 60/214,707
/ PRIOR FILING DATE: 2000-06-27
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 923
/ TYPE: PRN
/ ORGANISM: Homo sapiens
/ US-09-717-926-2

Query Match      25.0%; Score 597.5; DB 4; Length 923;
Best Local Similarity 30.8%; Pred. No. 6.9e-47;
Matches 152; Conservative 97; Mismatches 207; Indels 37; Gaps 13;

QY 5 YONYVNGEMKSSVN--QIEILSPIDDSIGFVPAMTREVVDHAKGREALP--AWAALT 60
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 443 YCCFINGQFTDADGKTVDITINPTDSTICKVSVALADVDKVALAKDAFENGEMGRN 502
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 61 VYERAOYLKADIIERDKEIATVLAKEISKAYNASV--TEVVRADLIRYAAE--EGIR 117
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 503 ABERGLIMYRLADILEENQEBELATIBALDSGAVYTLAKTHIGMSVQTRFYFAGWCDKIQ 562
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 118 LSTSDDEGKMDASTHKLAVIRROPVGIYLAIPVNYPNVLSGSKIAIPALIGSNVMMK 177
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 563 GSTIIPIN---QARPNRLFTTKEPLGVCAIIPNNYPMMLAMSAACLAAGNTLVLK 618
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 178 PPTGQSVSGVLVAKAPAEAGLPAGVFNITIGRSGEIGDYIVEHEVNFINTGSTPTVGOR 237
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 619 PAQVPLTLTKRELSELVAKGFPKGVNIIIPGSGIAGQLSHRPDIRKLGFTGISTPIGQ 678
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 238 IGR--LAKRPIMLELGGKAGIVLADADLONAAQIVAGAVDYSGORCTAIKRVLYVE 294
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 679 IMKSCAVSNLKKVSLGKSPILIFNDCELDRAVMGMGAVPFNKGNCIAAGRLFEVE 738
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 295 EVADELAEKISENVALLSVGDPEDNATVTPVID--DNGADFIESLVNDAKQKAKE--- 348
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 739 SIHDEFVTVTVEIKKIKGIDPLDRST---DHGPNKHALEKILQYCEFGVEGAT 792
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 349 ---LNEFRKRDGRLTPGLFDHYTLDMKLAMEEPFPIPIIRVK--DAEEAVAIANKSD 402
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 793 LVYCGRQVQRPGEFMEPTVFTDVEDMYLAKESFPFIWISKEFGDIDLDTLQRLNSTE 852
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 403 FGLQSVFTFRDFOKAFDIANKLEVGTVHIN--NKTGRGPDNPFPLGLKSGAGVQIGIRYS 460
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 853 YGLASGVFTFRDINKAMVSEKLEAGIVFINTYKTDVAA---PFGVKGSGFGKDLGEEA 909
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 461 IEAMNVKSIYVD 473
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 910 LNEYLTKTIVLE 922
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 59
US-09-902-540-15889
/ Sequence 15889, Application US/0902540
/ GENERAL INFORMATION:
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Wiegand, Roger C.
/ TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
/ FILE REFERENCE: 38-10(15849)B
/ CURRENT APPLICATION NUMBER: US/09/902,540
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: 60/217,883
/ PRIOR FILING DATE: 2000-07-10
/ NUMBER OF SEQ ID NOS: 16825
/ SEQ ID NO 15889
```

```
/ LENGTH: 477
/ TYPE: PRN
/ ORGANISM: Myxococcus xanthus
/ US-09-902-540-15889

Query Match      24.9%; Score 595.5; DB 4; Length 477;
Best Local Similarity 32.0%; Pred. No. 3.7e-47;
Matches 154; Conservative 86; Mismatches 224; Indels 17; Gaps 6;

QY 1 LTKEYQYVNGEMKSSVNOIILSPIDDSIGFVPAMTREVVDHAKGREALPMAALT 60
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 2 LAERYPYLANRPPQNALAVTHKYSGEVTHAVADAAVEEILAAVRAEPMKRLA 61
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 61 VYERAOYLKADIIERDKEIATVLAKEISKAYNASVTEVVRADLIRYAAEBSIRLST 120
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 62 PYARQAVLHECVRRREBAEELALALCTEAGKPLRDAGEVDRLDTTKAAAEAVR--- 118
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 121 SADBGRK-----DASTGHKLAVIRROPVGIYLAIPVNYPNVLSGSKIAIPALIGSNV 174
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 119 ---GGGEVLNLEVSRAAGYR--GFTQVVPVGPSPFTFPNPNLVAKHVAPALAAQCF 174
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 175 MEKPTGQSVSGVLVAKAPAEAGLPAGVFNITIGRSGEIGDYIVEHEVNFINTGSTPTV 234
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 175 ILKPSDRTPPVALLMAELIATELPEGAFSVLPVRLIEDIGP-LIEDRLKLSFTGSEKV 233
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 235 GQRIKLAGMPIMLELGGKAGIV--LADADLONAAQIVAGAVDYSGORCTAIKRVLY 292
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 234 GWELEARRAGRKRVVLELGSNAACVDEAPGAPLDTIARVAGQAFQAGQSCISQRLV 293
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 293 VEEVADLEAKISENVALLSVGDPEDNA--TVTVIDNSADFIESLVNDAKQKAKEINE 351
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 294 HASLYDALREKLIVERAKALRPGHPDEATTLGPMIDEPBARRLBEGMIOAAVORGARVL 353
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 352 FRDRLRLTPGLFDHYTLDMKLAMEEPFPIPIIRVDAEAVAIANKSDPGLQSVFT 411
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 354 GGRRSGLVDATVLEGVPPDASLSEEARGPVLLQFPDSFDEALAAQVNSGRRGLQAGLT 413
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 412 RDPQAFPIANKLEVTGTHINNKTGRGPDNPFPLGLKSGAGVQIGIRYSIEMTVKSIY 471
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 414 QLSAAMKAMBELEVGIVGVGVDPVSFRVDTMYPYGVSQSIGREBVKAIADMTLRLV 473
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 472 L 472
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 474 L 474
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 60
US-09-328-352-7337
/ Sequence 7337, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ PRIOR FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 7337
/ TYPE: PRN
/ ORGANISM: Acinetobacter baumannii
/ US-09-328-352-7337

Query Match      24.8%; Score 594; DB 4; Length 505;
Best Local Similarity 31.1%; Pred. No. 5.6e-47;
Matches 153; Conservative 109; Mismatches 198; Indels 32; Gaps 11;

QY 1 LTKEYQYVNGEMKSS--VNOIILSPIDDSIGFVPAMTREVVDHAKGREALP--AWA 57
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB 24 MARQGHFIDGKLVAEELDKVDIVNPSTEVVAQISIGSQDVSAVSAEHAQONMA 83
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 58 ALTVERAOYLKADIIERDKEIATVLAKEISKAYNAS--VTEVVRADLIRYAAEGCI 116
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
```

```
Db      84 ETTPIRYGVNTLADLIEQYGEILAQLETLSTGKLINIRHLEVAQSVFLTYFAGMAT 143
      117 RL-----STADEGKMDASTGHKLAVIRROPVGIATLAPYNPVNLGSKIAPALI 169
      144 KINGOTMOPISIPMOGEKTYAFT-----LRQPVGVAGIVPMNFSMLGWMKISALT 196
Qy      170 GGVNVMFKPTQOSVGLVLAKAFAGLPAQVFNITITGSGSEIGDYIVHEEVNFINFT 229
      197 TGCTTIVLKPEEPASLSLRLAELAIAGIPAGVINVTGKG-DTGQVLIESPLVKYVSFT 255
Qy      230 GSTPVGQRIKGLA---GMRPIMELGKDGAGIYLAADLDLNAKQIYAGAYDSGORCTA 286
      256 GSVPTGAIKGLKMLSSDLTIVSLELGKNAIYALADNIDELPTLIQTAFVHQGYCAS 315
Qy      287 IKKVLVEEYADELAEKISENVAKLSVDPEDNATV-TPVIDNSADFIESLVVDAROK- 344
      316 PERFFVHRTKYDELVDKLSKALSQFKISAMDEGSMGFLSNQPHFKVGHYLDMAKANN 375
Db      345 ----GAKELNEFRDGRLLTPGLFDHVTLDMKLAMEBPFGPIPIIRVKDAEAVAIANK 400
      376 QIIAGEEALDQI---GTFVQPTLISFRKNTDPLFSEBTFGPVVGVMPEFDEBLIQLMNQ 432
Qy      401 SDFGLQSVPTRFQKAFDIAFKLEVTGVIHNNKTKGRGPNPFLGKSGAGVQGIARS 460
      433 SRFGLTASTITNDLSKALRLIPKIEAGTLWVNNHTFLDP-SVPEGVKASGIGREPSDAP 491
Qy      461 IEAMTNVKSIVL 472
      492 IEDYTELKSVMI 503
Db
```

RESULT 61
US-09-252-991A-28449
; Sequence 28449, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28449
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28449

Query Match 24.7%; Score 591.5; DB 4; Length 748;
Best Local Similarity 30.3%; Pred. No. 1.8e-46;
Matches 148; Conservative 91; Mismatches 215; Indels 35; Gaps 7;
Qy 6 ONVYNGEMKSSVNOIEILSPIDSSLGFPVPMTRREVDHMKAGREALPAMALTYEYA 65
Db 265 KMINREBESKQVFNYPNATSELIGEVASGGAELDAVAARERFPMAATPAKERA 324
Qy 66 QYLAHKAADIIERDKEIATVLAKESKAYNASYTEVV-RTADLIRYAABEGIRLSTADE 124
Db 325 RLHRRIRGELIDRVNPHLALETDTGLPIHQTKVLIIPRASHNFEFFAIVCTRMN----- 379
Qy 125 GCGMDASTGHKLAVIR-----QPVGIYLAAPYNPVNLGSKIAAPLIGGNVNRK 177
Db 380 -----GHSYVDQOMLTNYLYQPVGCVLSPMNVNPPMTATWTATPLAIGNRAVAK 431
Qy 178 PPTQGSVSGVLAKAFAGLPAQVFNITITGSGSEIGDYIVHEEVNFINFTGSTVPGOR 237
Db 432 MSLSLPLTINELGRVLVHENGIPPGVINNVQGYGASAGDALVIRRDVRAVSFTGCTATGR 491

```
Qy      238 IGLAKMRPIMELGKDGAGIYLAADLDLNAKQIYAGAYDSGORCTAIKRVLYVEVA 297
      492 IMEAPGIRKYSMELGKSEVLVEFEDALRALDALFTFSLNGERCTAGSRIFVQESVY 551
Qy      298 DELAEKISENVAKLSVGDPEFNAT-----VTPVIDNSADFIE-----SLVVDAROK 345
      552 PQEVAEFAPARARLLVGDQDPRPTQVGSMTIOAHDKVTGYIRIGLEBAGTIVAGLEPR 611
Qy      346 AKELNEFKDGRLLTPGLFDHVTLDMKLAMEBPFGPIPIIRVKDAEAVAIANKSDFGL 405
      612 A-CLPAHLISKQGIQPTVFADVNRNRRIAQEBELFGVVCCLIPKDBAALRLANDVEYGL 670
Qy      406 QSSVPTRDQKAFDIANKLEVTGVIHNNKTKGRGPNPFLGKSGAGVQGIIRYSIEAMT 465
      671 ASYIWTQDIGKHLRLARGIEAGVVPINSQVNDLRO-PFGGVKASGTGREGREYSFEVRA 729
Qy      466 NVKSIYLDIM 474
      730 EIRNCTISM 738
Db
```

RESULT 62
US-09-328-352-6757
; Sequence 6757, Application US/09328352
; Patent No. 6562958

; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6757
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6757

Query Match 24.7%; Score 590.5; DB 4; Length 500;
Best Local Similarity 33.2%; Pred. No. 1.2e-46;
Matches 161; Conservative 76; Mismatches 225; Indels 23; Gaps 10;
Qy 1 LTKYQNYNGEMKSSVNOIEILSPIDSSLGFPVPMTRREVDHMKAGREAL--PAMAA 58
Db 12 MTAADFLYINGQFESGAATFESINPATGCVMAQMPFARTDEVNRAVQAASQLKASAWAG 71
Qy 59 LTVYERAOYLHKAADIIERDKEIATVLAKESKAYNASYTEVVRTADLIRYAABEGIRL 118
Db 72 LTAQGRKLYKLADLIERKAPQLAQIERTSDGKIIIRTSQIAVVAEYRYYA--GI-- 127
Qy 119 STSAD--EGKMDASTGHKLAVIRROPVGIYLAIPYNPVNLGSKIAAPLIGGNVNR 176
Db 128 ---ADLIESPFLIDKADQAMITVRBPVGVAIVAPMNSQLFLSAVKGPALAACTTVL 184
Qy 177 KPPTQGSVGLVLAKAFAGLPAQVFNITITGSGSEIGDYIVHEEVNFINFTGSTVPGQ 236
Db 185 KASBQPGGLLAFALVHENGFPAGVNVITTFGEGCAGVLSHDPDVHVAATGGETAR 244
Qy 237 RI--GKLAKMRPIMELGKDGAGIYLAADLDLNAKQIYAGAYDSGORCTAIKRVLYVE 294
Db 245 HIVRNSAENILAKYSLELGKSPFIYFADADIOSAVNAQVAAIFAAATGSCVAGSRLVEE 304
Qy 295 EVADELAEKISENVAKLSVGDPEFNAT--VTPVIDNSADFIESLVVDAROKAKEL---N 350
Db 305 SIKDEFQRLVLRVQTIKGLPHEMATERGPCTIAROKOKIQOVAASSIQGAKLLTGSK 364
Qy 351 EFKDGRLLTPGLFDHVTLDMKLA--WEPPGPIPIIRVKDAEAVAIANKSDFGLQS 407
Db 365 SLERAGYVYPITLD--CTDVPQADCIITELFGPILSVDSFQDEABAQVKANSTPGGLAA 422
Qy 408 SVPTDFQKAFDIANKLEVTGVIHNNKTKGRGPNPFLGKSGAGVQGIIRYSIEAMTV 467

Db 423 GFFTNLSRAHMTKVGISGIWLVNTYRAVSP-LAPFGGHGSHGREGAGNAVLDYTTT 481
 QY 468 KSIVL 472
 Db 482 KTWVL 486

RESULT 63

US-09-543-681A-4515
 ; Sequence 4515, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 4515
 ; LENGTH: 499
 ; TYPE: PRT
 ; ORGANISM: *Proteus mirabilis*
 US-09-543-681A-4515

Query Match 24.4%; Score 585; DB 4; Length 499;
 Best Local Similarity 28.4%; Pred. No. 3,9e-46;

Matches 139; Conservative 109; Mismatches 204; Indels 38; Gaps 9;

QY 8 YVNGEM---KSSVNOIEILSPIDDSLGFPAMTTEEDHAMKAGREALPMAALTVE 63
 Db 18 YHGGVVDCAQPECEBPFAINPATGEVIAHLQACEDINWAVESAKOGOKTMAAMPAA 77
 QY 64 RAOYHKAADIIRDKKEIATVLAKEISRAYNAS-VTEVVRTADLIRYAABEGIRLSTSA 122
 Db 78 RSRILRRADILIRERNDLAHLTETDTPGKPLSETRYVDIVTGADVLEYAA----- 127
 QY 123 DEGGKMDASTGKHLAV-----IRQPGIYALAPYVNVLSGSKTAPALIGNVYM 175
 Db 128 ---GLPAAEGRIPIRDSAFVYTRREPGLVAVGAGAMVPIQIALMKAPALAAGNAMI 184
 QY 176 FRPPTGVSGLVLAFAAAGLPAGVPTITGRGSEIDGYIHEHEVNFINTGSTPGV 235
 Db 185 FRPSEMTSLTKLALITYEAGLPQGVFNVTG-DATVQWMLQHDIAKSVSTGSDITG 243
 QY 236 QRI---GKLAKMRPIMLELGGKDAIVLADADLDNNAKOIVAGAYDYSQORCTAIRVLV 292
 Db 244 KKVMASSASSSLKDVMTLGGKSPLIIFDDADDDTAADIAMMANFYSSGQVCTNGTRFV 303
 QY 293 VEEVADLAEKISENVAKLSVGPDP-NATVTVVIDNSADPFIESLVNDAKQAKELNE 351
 Db 304 PEKLKQAFBAKIIERVSRIKRVGSLPMDNFGPLVSPHLEKLSYIEBKAGQAGALLCG 363
 QY 352 PKR-----DGLLTPGLFDHYTLPMKLAEEPPPIIPIRVKQAEVAALANKSDRG 404
 Db 364 GERLFAKALAKAYVAPYPTDCTDMKITDDEIFGPVMSISLYQDEYIIRANNSVYG 423
 QY 405 LOS\$VTRDFQAKAFDIANKLEVGTVAHNNKTGRGPNPFLGLKSGAGVQGIIRYSIEM 464
 Db 424 LAAGLVTRDITLTHRVYHQLHLEAGICWINT-WGESAEMEPVGYKHSVGRENGLVTLQNY 482
 QY 465 TYNKSTVLDLM 474
 Db 483 TQIKSTQVEL 492

RESULT 64
 US-09-134-001C-4246
 ; Sequence 4246, Application US/09134001C
 ; Patent No. 6380370

; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 4246
 ; LENGTH: 488
 ; TYPE: PRT
 ; ORGANISM: *Staphylococcus epidermidis*
 US-09-134-001C-4246

Query Match 24.4%; Score 583.5; DB 3; Length 488;
 Best Local Similarity 29.9%; Pred. No. 5.1e-46;

Matches 147; Conservative 95; Mismatches 214; Indels 35; Gaps 11;

QY 1 LTKEY-----QNVNGEMKSSVN-QIETLSPIDDSLGFPAMTTEEDHAMKAGRE 51
 Db 7 LRSEYNKRNFTKQYINGEWSDASGETTIDVINPTEVERMKIAGNEEDVNAKADADK 66
 QY 52 ALPAAALTYVERAQLHKAADIIRDKKEIATVLAKEISRAYNASVTEVVRTADLIRYA 111
 Db 67 VLEFRHSVBERRLDLKIKYKYNRKNDLIEATIDELGAPLSV-----ENVHY- 117
 QY 112 AEEGIRLSTAD--EGGKMDASTGKHLAVIRROPVGIYALAPYVNVLSGSKTAPAL 168
 Db 118 -QWGINHFTFAARDALDSFQFEORGDLLV--KEAIVAGLVTPWNPPTNQTSLTAAAF 174
 QY 169 ICGNVNMRPPTQGSVGLVLAFAAAGLPAGVPTITGRGSEIDGYIHEHEVNFINF 228
 Db 175 AAGSPVVLKPEEETFPFAITLAEIFDKVGPVGVNVLNKGSGGNPLSEHPKRVMSF 234
 QY 229 TGSSTVGORIGLAG--RPIMLLELGGKDAIVLADADLDNNAKOIVAGAYDYSQORCTA 286
 Db 235 TOSGPTGSKIMEKAKDKRYKSLHIGKSPYIVLDDVVEBAANAATTKKVANNCTQVCTA 294
 QY 287 IKRVLVEEVADELAEKISENVAKLSVGPDPNAT-VTVVIDNSADPFIESLVNDAKQK 345
 Db 295 GTRVILPESIKEDYLTVAKEAFSKVKGQPREBGTQVPIISKQPDQVQYIDKGINEG 354
 QY 346 AKEL-----NEPKDGLLTPGLFDHYTLPMKLAEEPPPIIPIRVKQAEVAALAN 399
 Db 355 AELFYGGPKPEGLDKGYFARPTTIFINVDMHTIAQEBIFGPVMSITYNNLDEAIELAN 414
 QY 400 KSDPGLQSSVFRDFQAKAFDIANKLEVGTVAHNNKTGRGPNPFLGLKSGAGVQGIIRY 459
 Db 415 DTKYGLAGVIGKDDTLRHVARSIEAGTIEL-NEAGKPD-LPFGYIKESGLGEMNDY 472
 QY 460 STEAMTVKSI 470
 Db 473 GIEEFLVYSI 483

RESULT 65
 US-09-489-039A-9023
 ; Sequence 9023, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342

QY 298 DELAEKISENVAKLSVGDEF-DNATVTPVIDNSADFIESLVVDARQKAKELN---BPK 353
DB 301 EALISOKLADIANAAQVGDGMASSTFEGPVQOMQYVKKALIAEALAQAKALSCQKLP 360
QY 354 RRGRLTLPGLFDHVTLMKLAWEERPGPIIPIIRVDAEAVAANKSPGIGSVTFRD 413
DB 361 EGGYIAPILITREVSSOSCHVOEBOGPVLPVKKTIDINDALAPANDSEFGIGSVTSSD 420
QY 414 FOKAFDIAKLEVGTVHINNKTGRGPDNPFLGLKSGAGVQIGRYSIEAM 464
DB 421 LKTAQTYATQLOCGTYVITWTHAEVLP-HAPFGWMKMSGLGAE---FGLEGL 467

RESULT 68
US-09-344-882-20
; Sequence 20, Application US/09344882
; Patent No. 6764851
; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J
; APPLICANT: Wurtzle, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinsan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allred, Carolyn C
; APPLICANT: Faltland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; TITLE OF INVENTION: Acetyl CoA Levels in Plants
; FILE REFERENCE: 201573
; CURRENT APPLICATION NUMBER: US/09/344,882
; CURRENT FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,717
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 20
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-09-344-882-20

Query Match 23.8%; Score 568.5; DB 4; Length 538;
Best Local Similarity 32.0%; Pred. No. 1.5e-44;
Matches 156; Conservative 85; Mismatches 208; Indels 39; Gaps 13;

QY 6 QNYVNGEWSKSVN--QIEILSPIDSSIGFVPAMTRREVDHAMKAGREALP--AMAAALT 61
DB 59 QLLINGNFVDSASGTFPTLDPRTGSEVIAHVAEGDEIDINRAVKAARTAFDEGPMWKSA 118
QY 62 VERAOYLKAAADIIRDKKEIATVLAKEISKAYNASVT--EVRTADLIYYA--AEGR 117
DB 119 YERSVTLRFADLVEHSEELASLETWNGKPYQOSLTAEIPMFALFFYVAGMDKTHG 178
QY 118 ILSADEGKMDASTGHKLAIVTRPOVGIIVLAIPYNPVNLSGSKIAPALIGANNVWPK 177
DB 179 LTIPLD--GNYOVHLH-----BPIGAGQIIPNPFILMFAWKVGPALACGNTIYVK 229
QY 178 PPTQSVSGVLAKAPAEAGLPAGVFNITITGSGEIGDIYVHEEVENFINTGSPVQGR 237
DB 230 TAEQPLPFAFYAGKLFLEAGLPFGVNLIVSGFATGALASHMDVDKLAFTGSDTDGKV 289
QY 238 ICKLA---GMRPIMLEIGKDGAGIYVLADDLNNAKQIYAGAYDVSQGRCTAIKRYLVE 294
DB 290 ILGLANSNLKEVTLLELGGKSPFIYFEDADIDKAVELAHFALFPNOGCCGAGSRFVHE 349
QY 295 EVADELAEKISENVAKLSVGDEPDNATVT--PVLD---DNSADFIESLVVD--ARQKAK 347
DB 350 KYIDFVEKSKRAALKRVVGDPEFRKIGEGFOIDLKQFEKWKYISGIESNAITTECGSD 409
QY 348 ELNEFRDGRLLTLPGLFDHVTLMKLAWEERPGPIIPIIRVDAEAVAANKSPDGLQS 407

DB 410 QIGD---KGYPLOPFVFSNVDDMLIADOLFEGPVQISILKSPSDVDEVKRNKRTYGLAA 466
QY 408 SVFTRDPKAFDIAKLEVGTVHINNKTGRGPDNP---PFLGJGSGAGVQIGRYSIEA 463
DB 467 GVFINLDTLAPRVSRALKAGTWNCF-----DVFDAAIPFGYKMSGNGREKGIYSLNN 521
QY 464 MTNVKSIY 471
DB 522 YLQIKAVV 529

RESULT 69
US-09-716-865-6
; Sequence 6, Application US/09716865
; Patent No. 6548292
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Micheal G
; APPLICANT: Nagarajan, Vasantha
; APPLICANT: Chen, Mario W.
; TITLE OF INVENTION: Bacterial Plasmid Having Genes Encoding Enzymes for the
; TITLE OF INVENTION: Degradation of Aromatic Compounds
; FILE REFERENCE: BC1016 US NA
; CURRENT APPLICATION NUMBER: US/09/716,865
; CURRENT FILING DATE: 2000-11-26
; PRIOR APPLICATION NUMBER: 60/167,062
; PRIOR FILING DATE: 1999-11-23
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Pseudomonas CT14
US-09-716-865-6

Query Match 23.5%; Score 562.5; DB 4; Length 486;
Best Local Similarity 30.3%; Pred. No. 4.8e-44;
Matches 149; Conservative 92; Mismatches 220; Indels 31; Gaps 10;

QY 3 KEYONYVNGEWSKSVN--QIEILSPIDSSIGFVPAMTRREVDHAMKAGREALP--WAA 59
DB 2 KDIRNPFINGEYVTNDSGTYEKRNKPNVDSLIGVHEAQCPYDAVAARALANGPWKLT 61
QY 60 TVYERAOYLH--KAADIIRDKKEIATVLAKEISKAYNASVTEVVRTADLIYYAAEGIRL 118
DB 62 SVYRCANLDDGVAAEINRRPDPFQAEIADGKRAHLASHIDIRGAANFKIPIDTIONV 121
QY 119 STSADGKMDASTGHKLAIVTRPOVGIIVLAIPYNPVNLSGSKIAPALIGANNVWPK 178
DB 122 STSEFEMKTPPGKTAARSYGV--RTPRGVIAIICPNLPLILMTWKCGRPAMACGNTVVVKP 179
QY 179 PPTQSVSGVLAKAPAEAGLPAGVFNITITGSG--SEIGDIYVHEEVENFINTGSPVQGR 237
DB 180 SEAPFTSTYTLLEGVNKKKGVPPGVYVNVNGVNSAGSFLTAHQGVNWTITTGELKITYA 239
QY 238 ICKLA---GMRPIMLEIGKDGAGIYVLADDLNNAKQIYAGAYDVSQGRCTAIKRYLVEE 295
DB 240 IMKAGADIRRVLSLELGGKMAAVPADCPFENALATVIRSAFENGQVCLGTERIYVERP 299
QY 296 VADLEAEKISENVAKLSVGDEPDNAT--VTPIVIDNSADFIESLVVDARQKAK----- 347
DB 300 IFDKVSALKERRAAIKRGRPFDDTKIGPLVSKIHQKVLISYAKAKABAGANIVLAGGV 359
QY 348 -ELNEFRDGRLLTLPGLFDHVTLMKLAWEERPGPIIPIIRVDAEAVAANKSPDGLQ 406
DB 360 PNMPPDLDDGGCWVEPTITGTPLESSPIREBEIIFGCGCHIQPDTREEEVLNNVNSPYLA 419
QY 407 SSVFTRDPKAFDIAKLEVGTVHINN-----KTGRGPDNPFLGLKSGAGVQIGRYS 460
DB 420 TSIHTQDISRAKRLATQLEVLGCMINSWFLADLRT-----PFGSGSQSGIGRGGIHS 472
QY 461 IEAMTNVKSIVL 472

Db 473 LEFYTELRNVI 484

RESULT 70

US-09-252-991A-32298
 ; Sequence 32298, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 32298
 ; LENGTH: 475
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-32298

Query Match 23.5%; Score 562; DB 4; Length 475;
 Best Local Similarity 33.3%; Pred. No. 5,1e-44;
 Matches 157; Conservative 81; Mismatches 195; Indels 38; Gaps 12;

21 EILSPIDSSLGFPVPMATREEDVHAMKAGREALP--AMALTYVERAOYLHKADIIERD 78
 15 DAIDPSNRLLARVAACDAADVAAARRAPDEGWARLAPVERKRVLLRLAEILAH 74
 79 KEERATV---LAKESKAYNASVTEVETADLIRYAEERGLSTSADEGKMDASTGH 134
 75 REELALLDLSLMGKRPWDAWN---IDVPGAHVFAWAE-----SLDLTYQVAPTAQ 124
 135 K-LAVIRROPVGIYLAIPNYVNVSGSKIAIPALIGANVMPKPTQSGVGLVAKAF 193
 125 QTLATITRVPLGVIYGVVPPNPLDMMAMKLAALAGNSVVLKPAEQSPFSLRLAEIL 184
 194 AEAAGLAVGAVNTTIGRSEIGDYIVHEEVNFINTGSTPVQ---RIGKLAMRPIIME 250
 185 LBAAGVPGVNAVPPGISEQKALGLHPEVDALVFTGSTVGRFMQYSAQSLMKQWLE 244
 251 LGGKDGIVLADA-DIDNAKQIVAGAYDYSQGRCTAIKRVLVVEEVADELAEKISENVA 309
 245 CGKSPNLVAFADCRDLDLALEKAFGLFFNQGSEVCANSRLVVERSTHDEFVERLLAKAR 304
 310 KLSVGPEDFNAT-VTPYIDNSADFIESLVVDARQKAKELNEFKDGRLLT----- 360
 305 DMQGPBLDPASBAGALVDRRQIAGILAAIERAQSEGAATLLG---CGRLTINSGDNFI 360
 361 -PGLFPHVTLDMKLAMEEPGPILPIIRVKDAEBAVAIAKSPFGLOSSVFTDFOKAP 419
 361 EPTLFGVDRDMQJLAREIFGPVLAISAPDSEBAIRLANDSRYGLASLSMSDLRRAHR 420
 420 IANKLEVTGVHINKTKRGPD-NFPFLGKSGAGVQIGIRYSIEAMTNVKS 469
 421 VARRIAGTVSV--TYDALDVAVPFGGKSGFGARDLSHSPDKTYQTKT 469

RESULT 71

US-09-252-991A-25398
 ; Sequence 25398, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 25398
 ; LENGTH: 642
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-25398

Query Match 23.4%; Score 561; DB 4; Length 642;
 Best Local Similarity 31.7%; Pred. No. 1e-43;
 Matches 150; Conservative 79; Mismatches 186; Indels 58; Gaps 12;

31 LGFVPMTRE---EVDHAMKAGREALPMAALTYVERAOYLHKADIIERDKERIATYL 86
 166 LGSTPLDADALMALDAVRAVDHQQGLWPNLRVARIQHYERFLARREGEAAVVKLL 225
 87 AKESKAYNASVTEVETADL-----RYAEERGLSTSADEGKMDAS 131
 226 MMEIGKLDKSKKEPDTCDYIVDTLEALKELDRSSRELEQG----- 269
 132 TGHKLAVIRROPVGIYLAIPNYVNVSGSKIAIPALIGANVMPKPTQSGVGLVAK 191
 270 ---TLQIRRVPLGVALCKGPNYVNVFTTILIPALIGANVMPKPAFVILLIRPLLE 326
 192 AFAEAGLAVGAVNTTIGRSEIGDYIVHEEVNFINTGSTPVQGRIGKLAKMRP---I 247
 327 APRDS-FPPGVIINVIGRRETVSALMASGKVDVFAFITHSGASDLKTL-HPRPHRLA 384
 248 MELGKDGIVLADADLNAKQIVAGAYDYSQGRCTAIKRVLVVEEVADELAEKISEN 307
 385 ALGLDKNGIYLPQVDLNAVNAEAVTGALSTFGRCCTALKLIFHESTVLRPFLDKFSAR 444
 308 VAKLSGDPED-NATYPIIDNSADFIESLVVDARQKAKELNEFKDGR--LTPGLP 364
 445 LSLKPGMEVGAALTPPEPKVDYVLEGLNDLAKARAVNAGGEGHRSFFYPALL 504
 365 DHVTLDMKLAMEEPGPILPIIRVKDAEBAVAIAKSPFGLOSSVFTDFOKAF 418
 505 SPVSPMRVYHEQGRGLPVVPYRHELDVIDYVLRSDVQGGSLSGNDPAQIGRLVDVF 564
 419 DIANKLEVTGVHINKTKRGPDNFPFLGKSGAGVQIGIRYSIEAMTNVKS 471
 565 --AN--QVRINVNAQCQRPDPSYFPMGRKNSABGTMSVHDALRVF-SRTIIV 612

RESULT 72

US-09-344-882-24
 ; Sequence 24, Application US/09344882
 ; Patent No. 6764851
 ; GENERAL INFORMATION:
 ; APPLICANT: Nikolau, Basil J
 ; APPLICANT: Wurtele, Eve S
 ; APPLICANT: Oliver, David J
 ; APPLICANT: Behal, Robert
 ; APPLICANT: Schnable, Patrick S
 ; APPLICANT: Ke, Jinsan
 ; APPLICANT: Johnson, Jerry L
 ; APPLICANT: Allred, Carolyn C
 ; APPLICANT: Fatland, Beth
 ; APPLICANT: Lutziger, Isabelle
 ; APPLICANT: Wen, Tsui-Jung
 ; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
 ; FILE REFERENCE: 201573
 ; CURRENT APPLICATION NUMBER: US/09/344,882
 ; CURRENT FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: US 60/090,717
 ; PRIOR FILING DATE: 1998-06-26
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: PatentIn Ver. 2.2

SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Arabidopsis Thaliana
US-09-344-882-24

Query Match 23.4%; Score 559.5; DB 4; Length 534;
Best Local Similarity 32.7%; Pred. No. 1.1e-43;
Matches 160; Conservative 76; Mismatches 220; Indels 33; Gaps 12;

QY 6 QNVVNGEMKSSV--QIETLSPIDSSLGFPVPMTRREVDHAKAGREALP--ANAALTV 61
DB 55 QLLIGRFVDAVSGKTFFLPDPNGEVLIAQVSEGDADVRAVAARAKAFDEGPPMKMA 114
QY 62 YERAOYLHKAADIIERDKEIATVLAKESKAYNASV--TEVVRADLIYYAAEEGRIST 120
DB 115 YERSKILRFADLIEGHNDIEALETWDMGKPRGSAQIEVPMIAVFVYIGMADKI-- 172
QY 121 SADEGGKMDASTGHKLAIVIRROPVGIYLAIAPYNYPNLSGSKIAPALIGNVVMEKPT 180
DB 173 ---HGWTMGDGRPHVQTL--HEPIGVAQGIIPWNPFLMLSMWLGDPALACGNTVVLKTA 228
QY 181 QSSVSGVLVAKAPAEAGLPAGVFNITTGSGSEIGDIYVHEEYNPINFGSTPPVGRIK 240
DB 229 QRPPLSALIVGKLHHEAGLPDGVNIVSGFATAGAAIASHMDVKVAFGSTDVCKIIE 288
QY 241 LA---GMRPIMELGKADAGIYLAADLDNAKQIVAGAYDVSQGRCTAIKRVLYVEEVA 297
DB 289 LASKSLKAVTIELESSEHSFPCEDADVQAVEIAHFAIPFNQGCCCGSGRTFVHERY 348
QY 298 DELAEKISENVAKLSVGDPPDNATVT--PVIDNSADPIESLV---VDA---RQKAKELN 350
DB 349 DEFVEKAKARALKRNVGDPFSGIEGQPVDSQPFKILKYIKHGEVAGTLQAGGDRLG 408
QY 351 EKRQDGRLLTPGLFPHVITLDMKLAMEEPPGPIPIIRVQDAEANAIAKSPGLQSSVF 410
DB 409 S---KGYVQPTVFSVDKMDMLATDEIFGPVQTLIKFKDDEVIARANNRSRGIAAGVF 465
QY 411 TRDFQKAPDIANKLEVGTVHNNKTGRGPD---NPPFLGKSGSAGVQIGRYSIRAMTN 466
DB 466 TQNLDTAHLMLRALRKGVTWVINC-----DVLDAIIPFGYKMGSGIGREKGIYSLANNYQ 520
QY 467 VKSIVLDMK 475
DB 521 VKAVVTSLK 529

RESULT 73
US-09-538-092-1173
Sequence 1173, Application US/09538092
Patent No. 675314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Manefield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CurataseqFormatter Version 0.9
SEQ ID NO 1173
LENGTH: 535
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number P51649
US-09-538-092-1173

Query Match 23.2%; Score 555; DB 4; Length 535;
Best Local Similarity 29.7%; Pred. No. 2.8e-43;
Matches 141; Conservative 97; Mismatches 220; Indels 16; Gaps 7;

QY 7 NYVNGEMKSSVNOIETLSPIDSSLGFPVPMTRREVDHAKAGREALPAMALTVERAQ 66
DB 63 SFVGGRWLPAAATFPVQDPASGAALGVADCGVEARAABAAYAFKRWREVSKESS 122
QY 67 YLHKAADIIERDKEIATVLAKESKAYNASVTEVVRADLIYYAAEEGRISTADSG 126
DB 123 LIRKNTMLIQKDDLARIITAESGPKLKHGELVLSAFLEWFSERARRV-----YGD 177
QY 127 KMDASTGHKLAIVIRROPVGIYLAIAPYNYPNLSGSKIAPALIGNVVMEKPTQGSVSG 186
DB 178 IIHFKAKORRLALVKQPIGVAVAVITPNNFPMITRKGAALLAAGCTIVVPAEDTPESA 237
QY 187 LVLAFAEAGLPAGVFNIT--TGRSGEIGDIYVHEEYNPINFGSTPPVGRIKLA- 242
DB 238 LALAEIASQAGIPSGVNVVIPCRRNAKEVEALCTDPLVSKISFTGTTGKILHHA 297
QY 243 -GMRPIMELGKADAGIYLAADLDNAKQIVAGAYDVSQGRCTAIKRVLYVEEVADELA 301
DB 298 NSVKVSMELGGLADPIFDSANVDQAVAGMASKFRRTGTCVCSNQFLVQRGTHDAFV 357
QY 302 EKISENVAK-LVSGDPPDNATVT--PVIDNSADPIESLVVDAROKAKELNEFKR---DG 356
DB 358 KAFAPAMKMLRVNGFPEGTTGGLINEKAVEKEKVNDVAVSGAIVVVGGRKHQLG 417
QY 357 RLITPGLFDHVTLDMKLAMEEPPGPIPIIRVQDAEANAIAKSPGLQSSVFTRDQK 416
DB 418 NFEFPTLLCNVTQDMLCTHEETFGPLAVIIFDEEEBAIAIANADVLGAGFYQDPAQ 477
QY 417 AFDINKLEVGTVHNNKTGRGPDNFPPLGKSGSAGVQIGRYSIRAMTNVXS 470
DB 478 IWRVAOLEVGVAGVNEGLISSV--CPFGVAKQSGLREGSKYGTIDEVLELKV 530

RESULT 74
US-09-328-352-7905
Sequence 7905, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
PRIOR FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7905
LENGTH: 564
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7905

Query Match 23.2%; Score 555; DB 4; Length 564;
Best Local Similarity 30.8%; Pred. No. 3.1e-43;
Matches 151; Conservative 93; Mismatches 208; Indels 38; Gaps 13;

QY 9 VNGE---W---KSSVNOIETLSPIDSSLGFPVPMTRREVDHAKAGREALP---P 54
DB 56 VDGELKMAAGATKTLSPITWIQPPGSIQOVELSGPYPMNGESEDALAAVRAVANNRG 115
QY 55 AMAALTVERAOYLHKAADIIERDKEIATVLAKESKAYNASVTEVVRADLIYYAAE 114
DB 116 EWPMAKVERIAQCMQNFQRMWEGRDLIYKIMEIGKSLADSEKEPRTITTYMR-----Q 171
QY 115 GIRLSTSADEGKMDASTGHKLAIVIRROPVGIYLAIAPYNYPNLSGSKIAPALIGNVV 174
DB 172 TIDALKDIDNANSRFVIAEGTIGQIRTRPLVGLQCMGPNYPLNTEFATLLIPALMGNTI 231
QY 175 MFKPPTQGSVSGVLVAKAPAEAGLPAGVFNITTGSGSEIGDIYVHEEYNPINFGSTPPV 234


```

Db      232  IFKPGQTLFPLLEAFRDS-FPKVINTIYAPGSLVPHLLASQINVLALLISSKV 290
QY      235  GORIGKLAKMRP-----IMLEIGKQDAGIVLADADIDNAKQIVAGANDYSQORCTAK 288
Db      291  ADHLKK---OHPKSHRLRALGLDAKVAAILLPDADIDLTVKECLGALSFNQORCTALK 347
QY      289  RLVVEBEVADELAKISENVAKLSVGDPPD-NATVTPVIDDNGADPFIESLVNARQKAK 347
Db      348  MLMVHRSIADDEFVNRITTEBLAKLVGPMWKGVSITPLPGMHRTAYMTTEIEDAVAKGAK 407
QY      348  ELN-----EFKRDRLITPGLFHDVTLDMKLAWEPEFPIITIRVKADEBAVAIANKSDP 403
Db      408  VVNPEGEFCK--TMFPAVAVVPTBGMRLYREGEQGPVPAVYDDIETVLDYVTTSDH 465
QY      404  GLQSSVFTRDFOKAFDIANKL--EVTGVHNNKTGRGPDNPPLGLKSGAGVQGIKRYSI 461
Db      466  GQGVSIQSDPAQIGHLVDTLVHQCVRVINCCQGRGPDVFPFGRKDSAGTLSVHDAL 525
QY      462  EAMTNVKSIV 471
Db      526  RAF-SIRSMI 534

```

RESULT 75

```

US-09-351-150A-9
; Sequence 9, Application US/09351150A
; Patent No. 6538177
; GENERAL INFORMATION:
; APPLICANT: Duvick, J
; TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification
; FILE REFERENCE: 035718/167676, 5718-62
; CURRENT APPLICATION NUMBER: US/09/351,150A
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Bacterium 2412.1
US-09-351-150A-9

```

Query Match 23.2%; Score 554.5; DB 4; Length 468;

Best Local Similarity 30.7%; Pred. No. 2,5e-43;

Matches 150; Conservative 91; Mismatches 205; Indels 43; Gaps 13;

```

QY      4  EYQNVYNGEKKSVNOIEILSPIDSSIGF-VPMATREEVDMHMKGREALPMAALTVY 62
Db      3  EYKLLIDGRIVAGATTMSVINPATETPLVIDCPRADRDQDEAVAAERAFQSMRATTLLE 62
QY      63  ERAQYLHKADIIERDEEIAITVLAKEISKAYNAVSTEVVRTADLIRYAEEGIRLSTSA 122
Db      63  QRATINAIADAEADQSALARLLTQEGKPLADAMGEIYASAAFRY-----FTSL 114
QY      123  DEGGKM--DASTGHKLAVIRQPVGYVLAPVNYPNLSGSKIAPALIGANNVMKPPPT 180
Db      115  DLPRVVRVDATGR--VEVHRRLPGVGCIVPWNFFMLMAFKIPALLAGNTVILKPPA 172
QY      181  QGSVSGVLAKAFABAGLPAGVFNITIGRSEIGDYVEHEEVNFNFGSTPVGORI-- 238
Db      173  TTPUTIALRFG-ALVKDITVPGVINIITD--ADDUGAENTRHGIRKISFTGSTQTKKQMA 230
QY      239  GKLAGMRPIMLEIGKQDAGIVLADADIDNAKQIVAGANDYSQORCTAIRVLVEEVAD 298
Db      231  GAAEGLRISLEIGANDALIVLDVDPEKAVAPRVFAAMQNAQOVCTIAAKRIYVHESLYE 290
QY      299  ELAEKISENVAKLSVGD-----PFDN----ATVPIVDNSADFIESLVVDARQK 345
Db      291  AMCEEFQQLARIVVGGLEQGVOMGPLQNRQPEKVLGLIERARD--GRITAGGRKG 348
QY      346  AKELNEKRDGRLLTTPGLFDHVTLDMLKLAWEPEFPIITIRVKADEBAVAIANKSDPGL 405
Db      349  DK-----GYFIEPIIVDIAEGALVDEEQGPVMPVIRYSDPYDAVRANASPYGL 400

```

```

QY      406  QSSVFTRDFOKAFDIANKLEVGTVHNNKTGRGPDNPPLGLKSGAGVQGIKRYSI EAMT 465
Db      401  GGSIMSRNVVRAMSLADMEAGSVWVKHADVQPD-LPLGAKFSQMSGLBEGGLHFT 459
QY      466  NVKSIYLDM 474
Db      460  QVQ--VLNM 466

```

Search completed: August 28, 2005, 10:29:22
Job time : 50 secs

This Page Blank (uspto)